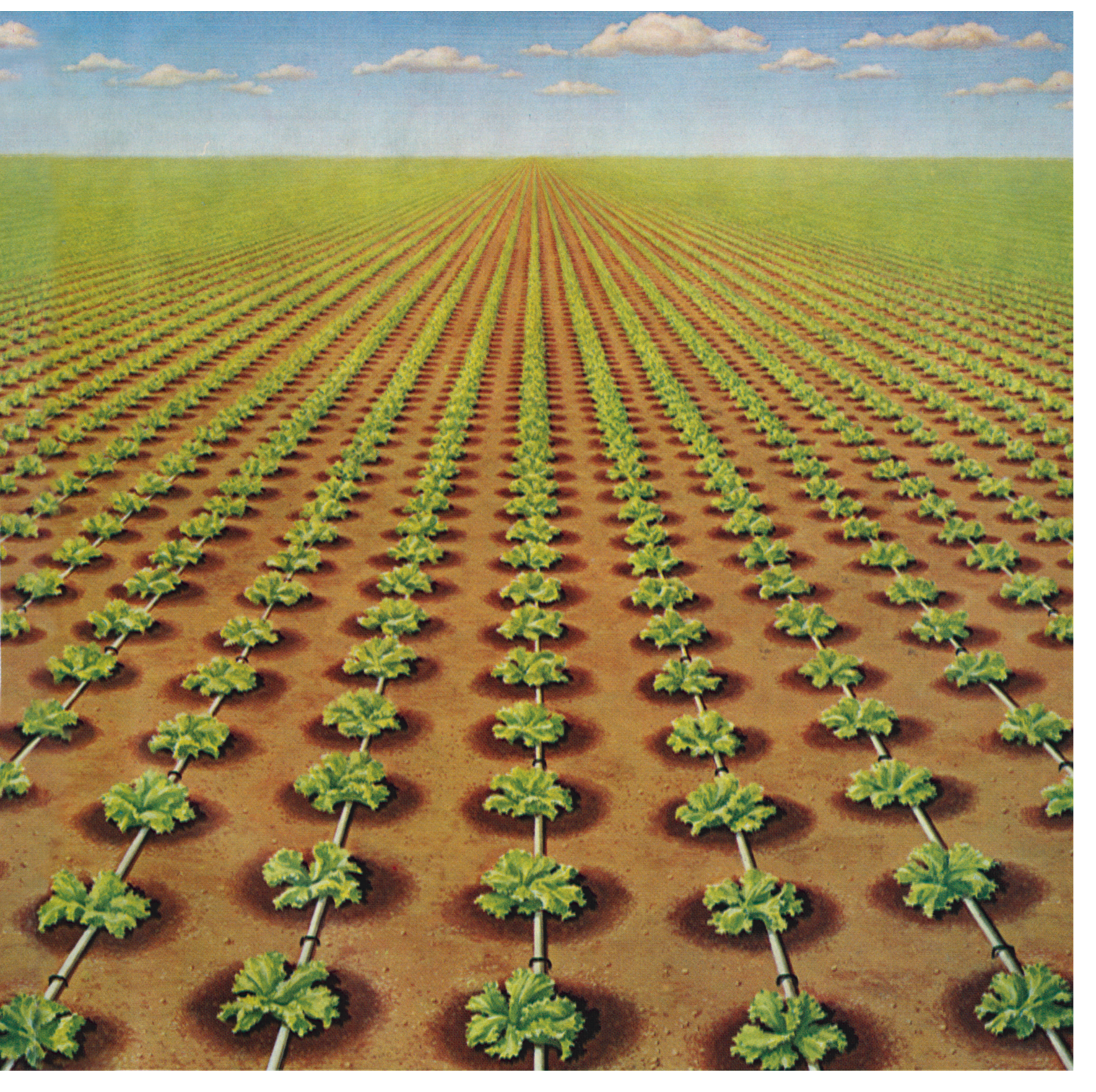
INVESTIGACION Y
CIENCIA

Edición en español de SCIENTIFIC AMERICAN



RIEGO GOTA A GOTA

Enero 1978

170 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 16
Enero 1978

INVESTIGACION Y
CIENCIA
Edición en español de **SCIENTIFIC AMERICAN**

EDITA Prensa Científica, S.A.
Calabria, 235-239
Barcelona-29
ESPAÑA

Colaboradores de este número:

Asesoramiento y traducción:

Manuel Puigcerver: *La búsqueda de vida en Marte;* M.ª Alba Fransí: *Riego gota a gota;* José Javier Labay: *El agrupamiento de las galaxias;* Monique Robert: *Gatos y vías comerciales;* Francisco Alférez: *Los usos de los instrumentos de sílex del Paleolítico;* Carles Oppenheimer: *El proceso de la fecundación;* Gerardo Espinosa: *Consecuencias de una anterior crisis energética;* Luis Bou: *Juegos matemáticos;* José E: Myro y Alicia Menéndez: *Taller y laboratorio.*

Ciencia y sociedad:

Pedro Miguel Echenique y Alfonso Balcells

Libros:
Philip Morrison



LA PORTADA

La pintura de la portada representa un campo de lechugas regadas con el sistemagota a gota, destinado a proporcionar, lenta y frecuentemente, las cantidades precisas de agua a la zona radicular de la planta. El agua se transporta a baja presión a través de tubos de plástico; se deposita en cada planta, literalmente casi gota a gota, por agujeros o emisores perforados (véase "Riego gota a gota", de Kobe Shoji, en este número). Comparado con otros métodos, el sistema reduce la tensión de las plantas y conserva el agua. La ausencia de maleza se debe en parte al sistema de riego gota a gota, ya que las zonas entre plantas y entre surcos están secas y la maleza no puede crecer.

Suscripciones:

Prensa Científica, S.A.
Departamento de suscripciones
Calabria, 235-239
Barcelona-29 (España)

Distribución para España:

Distribuciones de Enlace, S.A.
Ausias March, 49
Barcelona-10 (España)

Distribución para los restantes países:

Editorial Labor, S.A.
Calabria, 235-239
Barcelona-29 (España)

Precio de venta de este número:

España: 170 pesetas
Extranjero: 2 U.S. $

Condiciones de suscripción:

España:
Un año (12 números): 1870 pesetas
Extranjero:
Un año (12 números): 30 U.S. $

Ejemplar atrasado ordinario:
190 pesetas

Ejemplar atrasado extraordinario:
270 pesetas.

Solicitado control **OJD**

PROCEDENCIA DE LAS ILUSTRACIONES

Cubierta diseñada por Tom Prentiss

Página	Fuente
5	National Aeronautics and Space Administration
6	Dan Todd
7	U.S. Geological Survey
8-14	Dan Todd
17	I-Pai Wu, Universidad de Hawaii en Manoa
18-29	Alan D. Iselin
21	Type Systems Inc.
22	E. I. du Pont de Nemours & Company (*arriba*), Alan D. Iselin (*abajo*)
25	National Geographic Society-Palomar Observatory Sky Survey
26-27	Edward J. Groth, P. James E. Peebles, Michael Seldner y Raymond M. Soneira, Princeton University
28	George V. Kelvin
29	Edward J. Groth, P. James E. Peebles, Michael Seldner y Raymond M. Soneira, Princeton University
30-31	George V. Kelvin
32-34	Edward J. Groth P. James E. Peebles, Michael Seldner y Raymond M. Soneira, Princeton University
43-44	Enid Kotschnig
45-47	Bunji Tagawa
48	Type Systems Inc.
53	David Scharf

Página	Fuente
54	Lawrence H. Keeley
56-57	Tom Prentiss Lawrence H. Keeley
58	Tom Prentiss (*arriba*), Lawrence H. Keeley (*abajo*)
60	Type Systems Inc.
62	Mia Tegner, Scripps Institution of Oceanography
63	E. William Byrd
64	Carol Donner
65	Frank Collins, Stanford University
66-67	Carol Donner
68	Mia Tegner, Scripps Institution of Oceanography (*arriba izquierda y abajo*); Gerald P. Schatten y Daniel Mazia, Universidad de California en Berkeley (*arriba, derecha*)
69	Type Systems Inc.
70	Victor D. Vacquier, Universidad de California en Davis, cortesía de Academic Press, Inc.
75	John U. Nef
76-78	Dover Publications, Inc.
80	Ilil Arbel
84-91	Antoni Sellés
95-100	Ilil Arbel
102-105	Michael Goodman
106	Fritz Goro *(arriba),* Michael Goodman *(abajo)*

Dep. legal: B. 38.999-76
Fotocomposición: Fort, S.A.
Rosellón, 33 - Barcelona-29
Cayfosa. Santa Perpetua de Moguda. Barcelona.

Printed in Spain - Impreso en España

Los autores

NORMAN H. HOROWITZ ("La búsqueda de vida en Marte") es profesor y presidente de la División de Biología en el Instituto de Tecnología de California. Nacido en Pittsburgh, se educó en la Universidad de Pittsburgh y en el citado Instituto, de donde recibió el doctorado en Biología en 1939. Después de trabajar durante cuatro años en la Stanford University, regresó al Cal Tech (Inst. Tecnología de California) como profesor asociado de Biología en 1947. De 1965 a 1970 fue jefe de la Sección de Biociencias en el Laboratorio de Propulsión a Chorro del Cal Tech, y de 1970 a 1974 prestó servicio en la Comisión de Ciencias del Espacio de la Academia Nacional de Ciencias de los Estados Unidos.

KOBE SHOJI ("Riego gota a gota") es el vicepresidente de la Alexander & Baldwin Agribusiness Corporation en Honolulú. Se educó en el Pomona College y en la Universidad de California en Los Angeles, donde recibió su título de doctor en Fisiología Vegetal en 1950. Después pasó a la Facultad de la Universidad de Hawaii e investigó la nutrición mineral de las plantas de café, banano y caña de azúcar. Desde 1960 a 1975 trabajó para distintas compañías agrícolas de Hawaii, Irán y Puerto Rico, dirigiendo la investigación dedicada a hacer óptimas las producciones de la caña de azúcar y otros cultivos tropicales.

EDWARD J. GROTH, P. JAMES E. PEEBLES, MICHAEL SELDNER y RAYMOND M. SONEIRA ("El agrupamiento de las galaxias") son un grupo de cosmólogos de la Universidad de Princeton. Groth estudió en el Instituto de Tecnología de California y en Princeton, en donde obtuvo su doctorado en física en 1971. Se quedó allí y realizó estudios sobre los pulsars antes de emprender sus actuales investigaciones. Peebles es profesor de física en Princeton y jefe del grupo. Nacido en Canadá, asistió a la Universidad de Manitoba, trasladándose posteriormente a Princeton para conseguir su doctorado en física en 1961. Su interés inicial en la física gravitatoria le condujo a la astrofísica y, más tarde, a la cosmología. Seldner estudió en la Universidad Rutgers y ha conseguido este año su doctorado por Princeton. Su tesis doctoral versó sobre agrupaciones de radiofuentes. Soneira, que se halla preparando su doctorado, se licenció en física por la Universidad de Columbia en 1972.

NEIL B. TODD ("Gatos y vías comerciales") es director del Centro de Investigación sobre Genética de Carnívoros de Newtonville, Massachusetts. Estudió en las Universidades de Maryland, Massachusetts y Harvard, doctorándose en biología en esta última en 1963. De 1964 a 1968 estuvo contratado como genetista por el Centro de Investigación Animal de la Escuela de Medicina de Harvard, en el que se encargó de sendos programas de cruzamiento de perros y de gatos con fines de investigación. En 1968, entró a formar parte del Centro de Investigación de Genética de Carnívoros a la vez que trabajaba para varios laboratorios privados dedicados a la cría de animales y también ejercía como profesor adjunto de Biología de la Universidad de Boston, función que desempeñó hasta el año pasado. Los trabajos de Todd sobre genética de poblaciones de los gatos domésticos supusieron estudios de campo en todas las zonas del mundo.

LAWRENCE H. KEELEY ("Los usos de los instrumentos de sílex del Paleolítico") es antropólogo y tiene especial interés en las herramientas prehistóricas. Estudió en la Universidad de California en San José y se graduó en antropología por las Universidades de Oregón y de Oxford, recibiendo su doctorado en esta última la primavera pasada. Ha participado en investigaciones arqueológicas en California e Inglaterra y ha aplicado análisis por ordenador y microdesgaste a herramientas de sílex procedentes de yacimientos de Inglaterra, Alemania y Siria. "Por lo que respecta a investigaciones futuras, escribe, estoy interesado en aplicar mis métodos de investigación por microdesgaste a los instrumentos de piedra procedentes de Europa y Africa, pertenecientes al Paleolítico, y los del período Paleoindio de América del Norte. Trabajo también en la investigación de la tecnología preagrícola con especial referencia al uso de lascas de piedra como herramientas en la explotación de recursos vegetales."

DAVID EPEL ("El proceso de la fecundación") es profesor de biología en la Hopkins Marine Station, de la Stanford University. Empezó a interesarse por la biología siendo estudiante de la escuela superior, cuando le regalaron un viejo microscopio Leitz 1880, con el que pasaba largas horas observando protozoos. Continuó sus estudios de biología en la Universidad estatal de Wayne, hasta doctorarse en Zoología por la Universidad de California en Berkeley en 1963. Tras dos años de investigación postdoctoral en el laboratorio de Brillon Chance, de la Johnson Research Foundation de la Facultad de Medicina de la Universidad de Pennsylvania, se integró en 1965 en el grupo de trabajo de la Hopkins Marine Station. En 1970 se trasladó a la Scripps Institution of Oceanography de la Universidad de California en San Diego, pero retornó a Hopkins. Durante los tres últimos veranos ha codirigido el programa de embriología del Marine Biological Laboratory en Woods Hole, Massachusetts.

JOHN U. NEF ("Una anterior crisis energética") es profesor emérito de Historia Económica en la Universidad de Chicago. Nef nació en 1899, fue alumno del Harvard College y de la Robert Brookings Graduate School, en donde se doctoró en economía en 1927. Después de enseñar dos años en el Swarthmore College, pasó a integrar la facultad de la Universidad de Chicago. En 1942 fundó el Comité Pro Pensamiento Social, pequeña facultad interdisciplinaria dedicada al estudio del desarrollo histórico-cultural, que atrajo a muchos pensadores y artistas a Chicago. Entre éstos figuraron Friedrich von Hayek y Saul Bellow Es autor de 15 libros, incluyendo una historia de la industria hullera inglesa.

JESUS MOSTERIN ("La estructura de los conceptos científicos") cursó sus estudios en las Universidades de Madrid y Barcelona, doctorándose en filosofía por esta última. Durante 3 años estuvo investigando en el *Institut für mathematische Logik und Grundlagenforschung* (Instituto de lógica matemática e investigación de fundamentos) de la Universidad de Münster (Alemania). Actualmente es director del Departamento de Lógica y Metodología de la Ciencia de la Universidad de Barcelona. Es autor de dos libros de texto –*Teoría axiomática de conjuntos,* Barcelona 1971, y *Lógica de primer orden* (2.ª edición), Barcelona 1976– y de numerosos artículos técnicos sobre temas de lógica y filosofía de la ciencia. La idea de escribir este artículo le vino en julio de 1977, mientras daba un ciclo de conferencias sobre los conceptos científicos, como profesor invitado de la Universidad Nacional Autónoma de México.

La búsqueda de vida en Marte

Los vehículos Viking han completado sus experimentos biológicos. Estos experimentos no han descubierto procesos vitales, pero han revelado muchas cosas de interés sobre la química de la superficie del planeta

Norman H. Horowitz

La pregunta "¿Existe vida en Marte?" es interesante y científicamente legítima, con independencia de que generaciones de escritores de fantasía científica hayan poblado ese planeta con criaturas producto de su imaginación. De todos los cuerpos extraterrestres del sistema solar, Marte es el más parecido a la Tierra y, con gran diferencia, el hábitat más plausible para la vida extraterrestre en dicho sistema. Por esa razón, uno de los objetivos primordiales de la misión Viking a Marte era la búsqueda de indicios de vida.

Los dos vehículos espaciales Viking se lanzaron desde Cabo Cañaveral en el verano de 1975. Cada uno constaba de un módulo orbital unido a otro capaz de posarse. Cuando los vehículos llegaron a Marte, en julio y agosto de 1976, se puso a cada uno en una órbita predeterminada alrededor del planeta, y comenzó entonces la búsqueda de lugar apropiado para el descenso. La principal fuente de información sobre la que basar la elección del lugar de descenso eran las cámaras montadas en los módulos orbitales; también los sensores de infrarrojo que llevaban dichos módulos y los observatorios de radar en la Tierra suministraron importantes datos. La única consideración en la selección final de los lugares fue la seguridad de la nave. Sería un error suponer, sin embargo, que dichos lugares carecían por ello de interés biológico. Los criterios biológicos dominaron en las decisiones iniciales referentes a la latitud a la que se posaría cada nave espacial. Una vez determinadas las latitudes, existía relativamente poca diferencia entre unos lugares y otros a distintas longitudes.

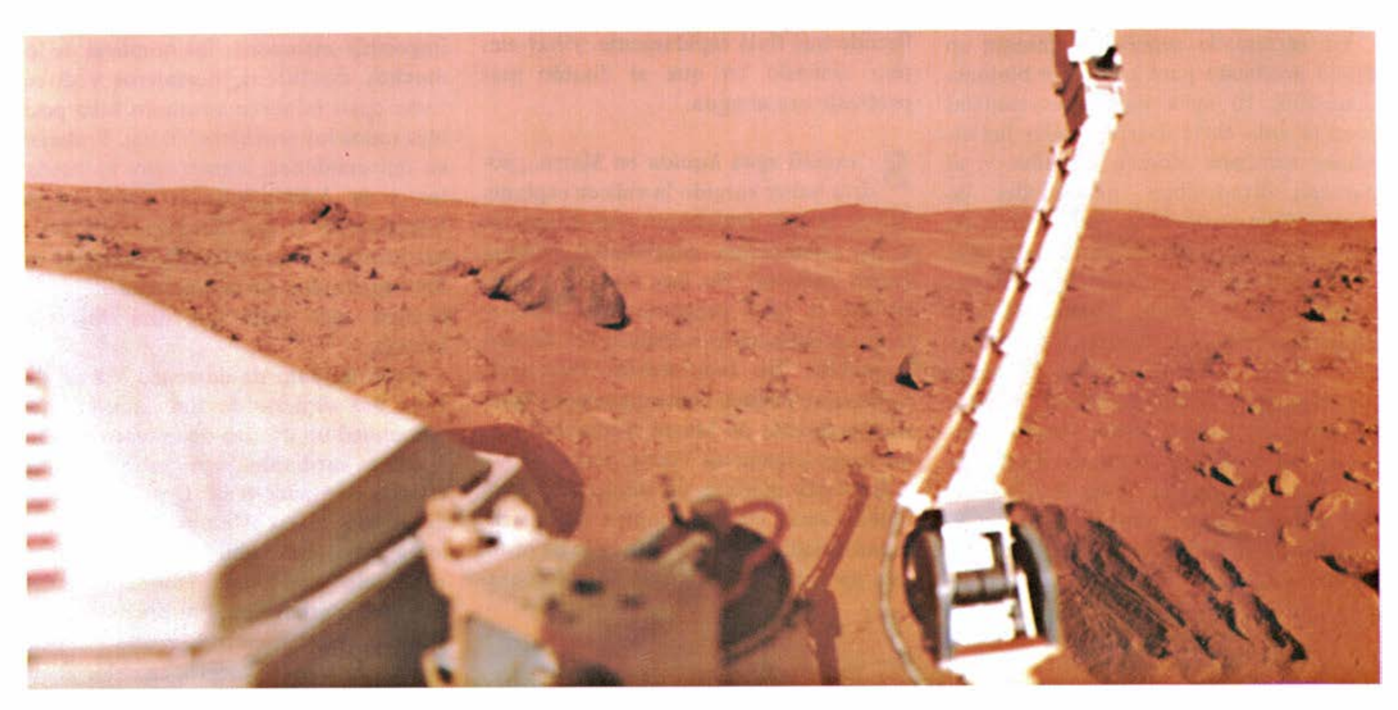
A una orden transmitida desde la Tierra, cada módulo de descenso se separó de su correspondiente orbital. Con la ayuda de sus retrocohetes y paracaídas, descendió hasta la superficie de Marte. Los dos módulos orbitales continuaron alrededor del planeta, haciendo funcionar sus propios instrumentos científicos y retransmitiendo a la Tierra datos emitidos por los de descenso. Las dos tomas de contacto se produjeron en el hemisferio norte de Marte en el verano marciano. (Marte tiene una sucesión de estaciones semejante a la de la Tierra, si bien cada estación dura casi el doble. El año marciano tiene 687 días marcianos; cada día marciano, llamado "un sol" por el grupo Viking para distinguirlo del día terrestre, tiene 24 horas y 39 minutos.) El día 20 de julio de 1976, el módulo de descenso *Viking 1* se posó en la región de la Planitia Chryse de Marte, a unos 23 grados al norte del ecuador. Seis semanas después, el módulo *Viking 2* se posó en la región de la Planitia Utopia, a unos 48 grados al norte del ecuador. En cuanto a longitud, los dos módulos están separados por casi exactamente 180 grados, lo que los coloca en caras opuestas del planeta. Como la instrumentación de ambos módulos es idéntica, la única distinción reside en la diferencia entre sus lugares de descenso.

La primera tarea biológicamente importante llevada a cabo por cada uno de los módulos de descenso fue el análisis de la atmósfera marciana. La vida se basa en la química de elementos ligeros, en particular, del carbono, hidrógeno, oxígeno y nitrógeno. Para ser adecuado como morada de vida, un planeta debe tener esos elementos en su atmósfera. Las observaciones espectroscópicas, hechas desde la Tierra y desde vehículos espaciales que habían pasado cerca de Marte en años anteriores, habían demostrado ya que el dióxido de carbono era el principal componente de la atmósfera marciana. Se habían detectado también pequeñas cantidades de monóxido de carbono, oxígeno y vapor de agua. No se había detectado, sin embargo, forma alguna de nitrógeno, y la teoría atmosférica sugería que Marte había perdido la mayor parte de su nitrógeno en el pasado.

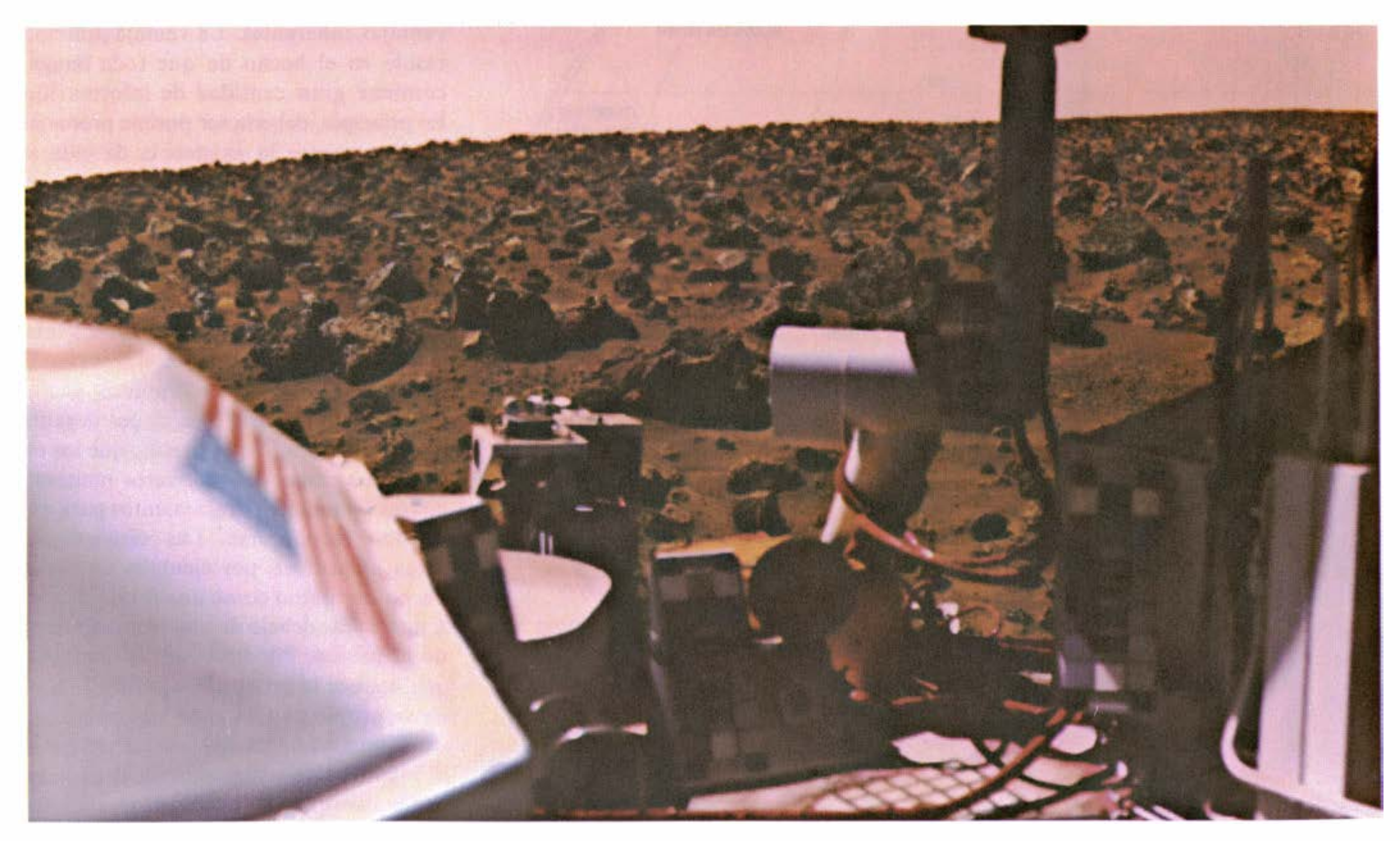
Cada módulo Viking de descenso analizó la atmósfera por medio de dos espectrómetros de masas. Un espectrómetro, que operaba durante el descenso hacia la superficie, tomaba muestras y analizaba los gases atmosféricos cada cinco segundos. El segundo espectrómetro operaba en el suelo. Los resultados mostraron que la atmósfera en la vecindad del suelo estaba constituida por aproximadamente 95 por ciento de dióxido de carbono, 2,5 por ciento de nitrógeno y 1,5 por ciento de argón, y que también contenía trazas de oxígeno, monóxido de carbono, neón, kriptón y xenón. En ambos lugares de descenso, la presión atmosférica era de 7,5 milibares. (La presión atmosférica al nivel del mar en la Tierra es de 1013 milibares.)

Como los vehículos espaciales Viking revelaron que el nitrógeno está realmente presente en la atmósfera de Marte, podemos decir que el planeta dispone de los elementos necesarios para la vida. Hay, sin embargo, un compuesto críticamente importante que no está en la lista de gases: el vapor de agua. Aunque las mediciones anteriores habían mostrado la existencia de trazas de vapor de agua en la atmósfera marciana, la cantidad varía con la estación y el lugar. Los módulos orbitales Viking llevaron a cabo un censo del vapor de agua sobre la totalidad del planeta con espectrómetros infrarrojos. Los resultados mostraron que la máxima concentración de vapor de agua atmosférico se daba en el borde del casquete polar septentrional (hemisferio de verano), y que la concentración descendía hacia el sur (lo contrario de lo que se encuentra en la Tierra). En la región polar, la cantidad de vapor de agua en la atmósfera formaría una película de sólo una décima de milímetro de espesor si todo él se condensara sobre la superficie del planeta. En los lugares de descenso, la concentración de vapor de agua oscilaba entre 10 y 30 por ciento de la concentración en el polo.

Estas cifras establecen en términos cuantitativos un hecho conocido desde hace largo tiempo: Marte es un lugar muy seco. Marte tiene hielo en sus polos, pero en ninguna parte de su superficie hay océanos, lagos u otras superficies de agua líquida. La ausencia de agua líquida está relacionada con la sequedad de la atmósfera por medio de una ley fundamental de la química física: la regla de las fases. La regla de las fases afirma que, para que pueda existir agua líquida en la superficie de un planeta, la presión del vapor de agua en la atmósfera debe en algunos períodos y en algunos lugares valer al menos 6,1 milibares. Las mediciones obtenidas por los Viking indican que la presión del vapor de agua en la superficie de Marte en el hemisferio norte es, a lo más, de 0,05 milibares aun en el caso de que todo el vapor de agua



LUGAR DONDE SE POSO el módulo de descenso *Viking 1* en Marte, fotografiado desde la nave espacial, en febrero, para poner de manifiesto la excavación de la serie de trincheras que se ven en la parte baja derecha. Las muestras de suelo tomadas de las trincheras se colocaron en los instrumentos montados en el vehículo espacial para hacer ensayos sobre su composición química y posibles señales de vida. El tinte naranja que recubre todo el cielo demuestra cromáticamente la existencia de una gran cantidad de polvo que está suspendido en la atmósfera marciana.



LUGAR DONDE SE POSO el módulo de descenso *Viking 2;* es un campo de piedras, superficialmente similar al sitio en donde se posó el *Viking 1.* La cámara a bordo del vehículo sacó esta fotografía a primeras horas de la tarde del 24 de septiembre de 1976, tres semanas después de que el vehículo se posara y antes de que se tomaran las muestras de suelo. El horizonte aparece inclinado porque el vehículo no se halla completamente nivelado. Los objetos que aparecen en primer término son algunos instrumentos del equipo transportados a bordo del vehículo.

estuviera concentrado en la baja atmósfera. A una presión tan baja, el agua no puede mantenerse en fase líquida; en razón de la temperatura, se congelará o se evaporará. Por la misma causa, no pueden formarse gotas de agua de lluvia en la atmósfera marciana y no puede fundirse hielo en la superficie de Marte.

La extremada sequedad presenta un difícil problema para cualquier biología marciana. El agua líquida es esencial para la vida en la Tierra. Todas las especies terrestres poseen elevadas y, al parecer, irreducibles necesidades de agua: ninguna de ellas podría vivir en Marte. Si existe vida en Marte, debe operar de acuerdo con un principio diferente por lo que se refiere al agua. Si Marte hubiera poseído un medio más favorable en el pasado, sin embargo, y si el planeta no se hubiera secado demasiado pronto, las especies podrían haber tenido tiempo de evolucionar y adaptarse a las condiciones actuales. Las fotografías realizadas por el vehículo espacial *Mariner 9,* que estuvo en órbita alrededor de Marte en 1971, sugerían que Marte podía realmente haber tenido agua corriente en su superficie en épocas anteriores. Las fotografías de los módulos orbitales Viking confirmaron esa impresión. La prueba la suministran los canales en el desierto marciano, que semejan lechos secos de ríos. No hay excesivos reparos en admitir que los canales fueron excavados por líquido que fluía rápidamente, y hay amplio acuerdo en que el líquido más probable era el agua.

Si existió agua líquida en Marte, ¿podría haber surgido la vida en el planeta? Si la vida evolucionó para adaptarse a las condiciones cambiantes, ¿podría existir todavía? No hay manera de responder a tales preguntas por razonamientos deductivos o incluso por experimentación en laboratorios terrestres. Solamente pueden contestarse por exploración directa de Marte, y eso es lo que las naves espaciales Viking hicieron.

En cada módulo de descenso Viking había cinco clases distintas de instrumentos dedicados a la búsqueda de pruebas de vida: dos cámaras para fotografiar el paisaje, un cromatógrafo de gases combinado con espectrómetro de masas para analizar la superficie en busca de materia orgánica y tres instrumentos proyectados para detectar las actividades metabólicas de cualesquiera microorganismos que pudieran estar presentes en el suelo. En esta breve reseña me será imposible mencionar los nombres de los muchos científicos, ingenieros y directivos cuyo esfuerzo conjunto hizo posibles todos los proyectos Viking. Trabajan en universidades, laboratorios industriales y la Administración Nacional de Aeronáutica y del Espacio (NASA) y sus campos de pruebas. Sus nombres están registrados en la creciente literatura técnica que trata de esta histórica misión.

Cada módulo de descenso Viking llevaba dos cámaras de tipo facsímil, que componen un cuadro de la escena explorándola mediante una serie de tiras estrechas y sucesivas. La construcción de la imagen en este tipo de cámaras es lenta, pero son robustas y de múltiples usos. Su resolución, moderadamente alta, alcanzaba varios milímetros a una distancia de 1,5 metros. Producían fotografías en blanco y negro, en color y estereoimágenes. Entre las dos cámaras de cada módulo podían explorar la totalidad del horizonte en torno al vehículo.

Como método de detección de vida, las cámaras tienen ventajas y desventajas inherentes. La ventaja principal reside en el hecho de que toda imagen contiene gran cantidad de información. En principio, debería ser posible probar inequívocamente la existencia de vida en Marte con una sola fotografía. Por ejemplo, si se viera una hilera de árboles en el horizonte, o si una mañana apareciesen huellas de pies en el suelo delante del módulo, no cabría duda alguna de la existencia de vida en Marte. Otra ventaja consiste en que la prueba impresa es independiente de toda hipótesis acerca de la química y fisiología de los organismos marcianos. No es preciso que los organismos respondan de cierta manera a ciertas sustancias o tratamientos para que puedan reconocerse. Las cámaras podrían identificar, por ejemplo, un hongo hecho de titanio como una forma de vida si apareciese debajo de una roca en el curso de la misión. Por supuesto, la confianza que merece la prueba fotográfica se basa en su propio conjunto de hipótesis acerca de la morfología de los seres vivientes. El inconveniente más claro de la cámara, como instrumento detector de vida, estriba en que puede existir todo un mundo de vida por debajo de los límites de resolución de la cámara.

De todos los resultados de la misión

LAS POSICIONES DE LOS DOS MODULOS de descenso Viking aparecen indicadas en este mapa de Marte, que muestra algunos de los rasgos geológicos más importantes del planeta. Los dos vehículos espaciales están en las caras opuestas, separados por unos 7400 kilómetros. Ambos están en el hemisferio norte, en lugares elegidos, en parte, por su posible interés biológico. El modulo *Viking 1* se encuentra en la región de Chryse Planitia, a una latitud de 23 grados al norte del ecuador; el *Viking 2* se halla en la región de Utopia Planitia, a una latitud de 48 grados.

Viking, las maravillosas fotografías del desierto marciano en los dos lugares de descenso son los más impresionantes. Las fotografías han sido cuidadosamente exploradas por ojos avizores y esperanzados, pero ningún investigador ha descubierto todavía nada que sugiera una forma viviente.

El paso siguiente fue analizar el suelo en busca de algún componente orgánico. Entre los elementos, el carbono es único en el número, variedad y complejidad de compuestos que puede formar. Las especiales propiedades del carbono, que lo hacen capaz de formar grandes y complejas moléculas, residen en la estructura básica del átomo de carbono. Esta estructura capacita al átomo de carbono para formar cuatro fuertes enlaces con otros átomos, incluyendo otros átomos de carbono. Las moléculas así formadas son muy estables a temperaturas ordinarias; tan estables que no parece haber límite para el tamaño que pueden alcanzar. La relación entre vida y química orgánica (química del carbono) reside en el hecho de que los atributos mediante los cuales identificamos seres vivientes –su capacidad para producir réplicas de sí mismos, para repararse a sí mismos, para evolucionar y adaptarse– se originan en propiedades que son exclusivas de grandes moléculas orgánicas. Son las proteínas y ácidos nucleicos, altamente complicados y ricos en información, los que dotan a los seres vivientes que conocemos, incluso los más simples como las bacterias y los virus, de su naturaleza esencial. Ningún otro elemento, incluyendo el favorito de los escritores de fantasía científica, el silicio, tiene la capacidad del carbono para formar esas grandes y complejas estructuras que son tan estables. No se debe a la casualidad que, aun cuando el silicio sea mucho más abundante que el carbono en la Tierra, sólo desempeñe papeles secundarios y no esenciales en bioquímica. La bioquímica es en gran proporción una química del carbono.

Tales hechos fundamentales conducen a la conclusión de que dondequiera que surja la vida en el universo, estará con toda probabilidad basada en la química del carbono. Este punto de vista ha sido fortalecido por el descubrimiento de compuestos orgánicos de interés biológico en meteoritos y en nubes de polvo, en el espacio interestelar. Aunque tales compuestos tienen un origen no biológico, se hallan estrechamente relacionados con los aminoácidos y nucleótidos, que son los materiales de construcción de las proteínas y de los ácidos nucleicos, respectivamente. El hecho de que se hayan formado en lugares alejados de la Tierra implica que la química del carbono da lugar, en cualquier región del universo, a compuestos conocidos. Este dato sugiere a su vez que la vida en otras partes del universo debe basarse en una química orgánica semejante a la nuestra, aunque no necesariamente idéntica.

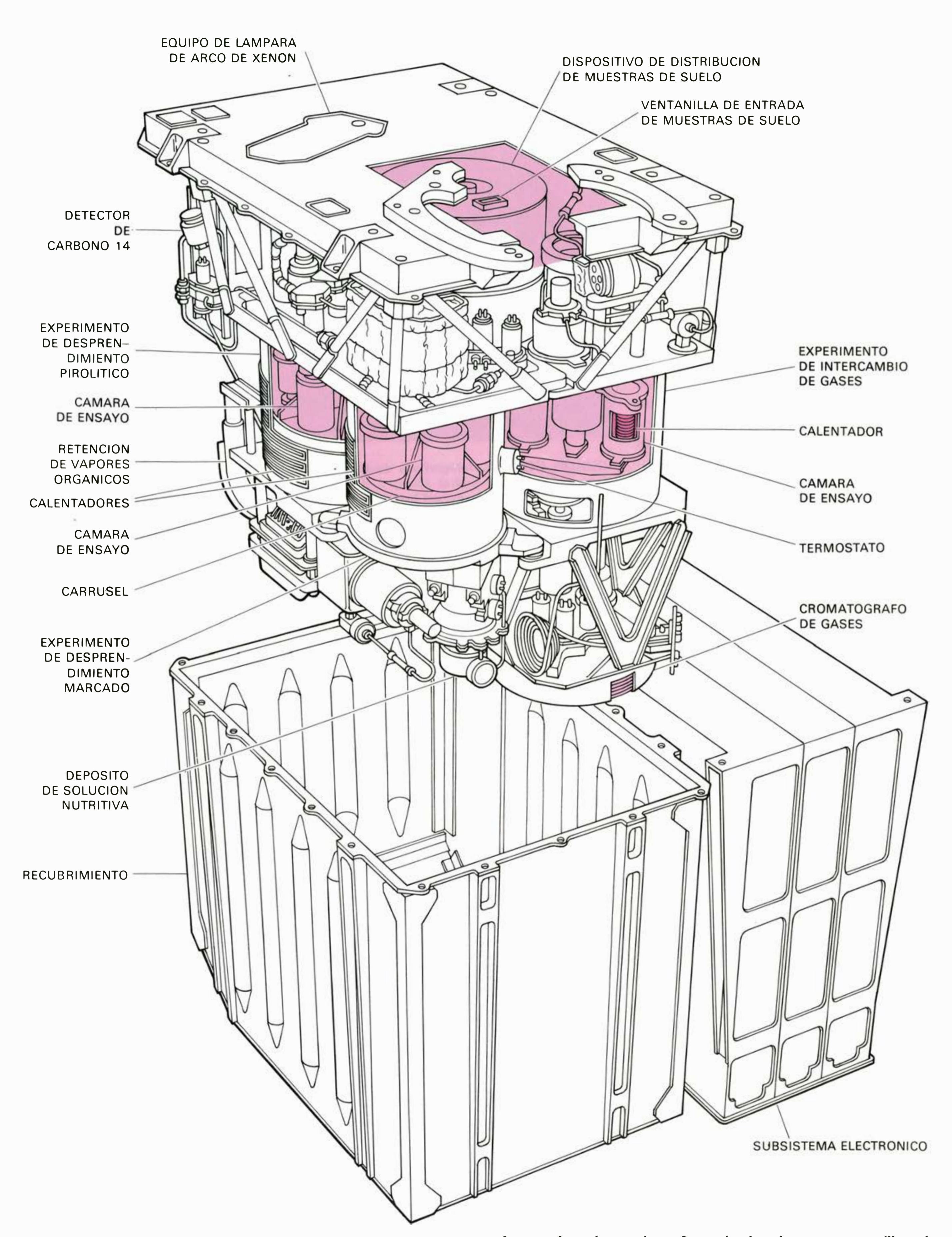
Tal consideración condujo a la decisión de incluir un experimento de análisis orgánico a bordo de los módulos de descenso Viking. El instrumento usado en el experimento fue el espectrómetro de masas que había analizado la atmósfera, combinado con un cromatógrafo de gases y un horno de pirolisis. Primeramente, una muestra de suelo marciano se calentó en el horno en diversas etapas hasta una temperatura de 500 grados centígrados. Se hizo pasar a las materias volátiles desprendidas por el cromatógrafo de gases. Como cada uno de los



REGION CHRYSE, fotografiada desde una altura de 1555 kilómetros sobre la superficie marciana por el módulo orbital *Viking 1* el 17 de julio de 1976. Muestra una posición muy exacta, recientemente determinada, del lugar de toma del módulo de descenso *Viking 1 (cruz).* La zona de toma, en la parte occidental de Chryse Planitia, es una suave llanura con numerosos cráteres pequeños de impacto salpicando la superficie. Las arrugas que se divisan hacia el oeste *(parte superior)* parecen similares a crestas volcánicas halladas en los suelos de lava lisos de los mares de la Luna.



REGION DE UTOPIA, fotografiada desde 3360 kilómetros sobre la superficie de Marte por el vehículo orbital *Viking 2* el 16 de agosto de 1976; la nave se hallaba inspeccionando el planeta en busca de un lugar para la toma de posición del módulo de descenso *Viking 2.* La región es escarpada, recubierta según parece por dunas. El lugar donde se posó el módulo de descenso *Viking 2* está indicado por una cruz en la parte situada más a la derecha (norte) de esta fotografía.



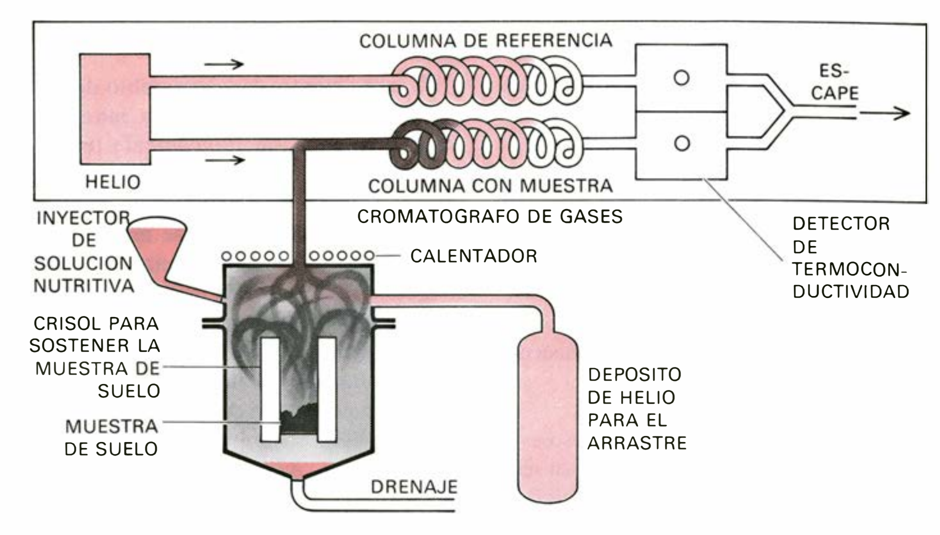
LABORATORIO BIOLOGICO a bordo de ambos vehículos espaciales Viking. Ocupa un volumen de sólo 28 decímetros cúbicos. Los tres experimentos biológicos fueron el de intercambio de gases *(derecha)*, el experimento de desprendimiento marcado *(izquierda, abajo)* y el experimento de desprendimiento pirolítico *(izquierda, arriba)*. Cada experimento, mostrado en corte aparte, tenía diversas cámaras de prueba en un carrusel, de modo que el experimento podía ensayar muestras diferentes de suelo marciano. Se vertía el suelo por una trampilla en la parte superior del laboratorio, donde caía en el interior de una tolva. Para cada muestra de suelo se hacía girar una cámara de prueba de cada experimento hasta quedar bajo la tolva para recibir una porción de la muestra. En total, fue una media docena de muestras de suelo las ensayadas en cada lugar de descenso. Los experimentos quedaron completados en abril. Los resultados se dan en las páginas siguientes.

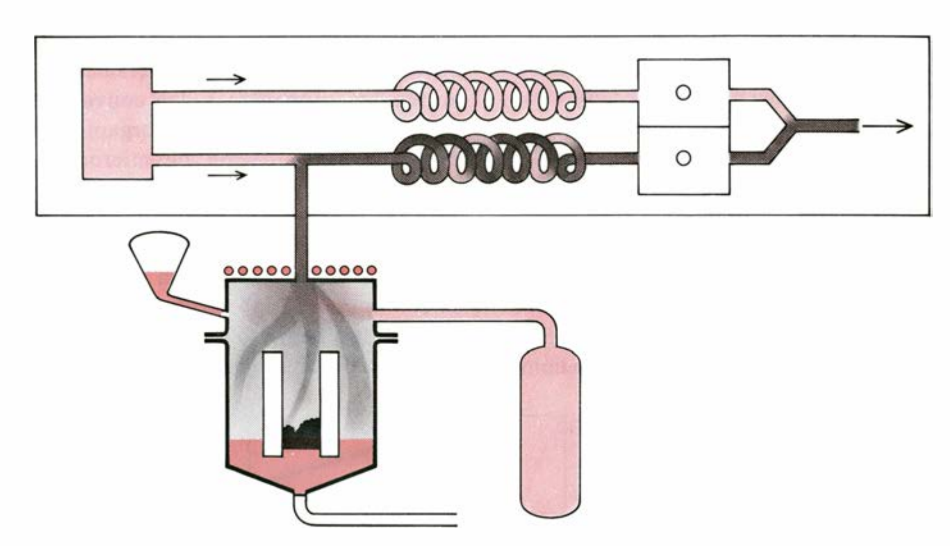
distintos compuestos tiene peso molecular, composición y polaridad diferentes, además de otras propiedades, pasa a través de las columnas del cromatógrafo de gases a una velocidad propia, de modo que los compuestos se van separando unos de otros. A medida que cada compuesto emergía de la columna cromatográfica, se dirigía al espectrómetro de masas para su identificación. Puesto que toda materia orgánica se rompe, o descompone en fragmentos más pequeños, a 500 grados Celsius, el método puede detectar compuestos orgánicos que tienen un amplio intervalo de pesos moleculares.

Se analizaron dos muestras de suelo en cada lugar de descenso. Los únicos compuestos orgánicos detectados fueron trazas de disolventes de limpieza que se sabía habían estado presentes en los aparatos. El hecho de que se detectaran los disolventes sirvió para demostrar que los instrumentos funcionaban correctamente. Las muestras calentadas desprendieron dióxido de carbono y una pequeña cantidad de vapor de agua; no se encontró nada más.

Este resultado es sorprendente y pesa poderosamente contra la existencia de procesos biológicos en Marte. La combinación de cromatógrafo de gases y espectrómetro de masas a bordo de cada módulo de descenso Viking es un instrumento sensible, capaz de detectar compuestos orgánicos en concentraciones de unas pocas partes por mil millones, nivel que está entre 100 y 1000 veces por debajo de su concentración en suelos de desierto en la Tierra. Aun cuando no haya vida en Marte, se había supuesto que la caída de meteoritos en la superficie marciana habría aportado suficiente materia orgánica al planeta como para que se la hubiese detectado. Por estar Marte cerca del cinturón de asteroides, del cual se originan los meteoritos, se cree que recibe un número de impactos de meteoritos mucho mayor que la Tierra o la Luna. Un problema que se discutió frecuentemente antes de que se lanzaran las naves espaciales Viking era si sería o no posible distinguir la materia orgánica biológica en Marte de la materia orgánica meteorítica que se esperaba estuviese presente. La ausencia de materia orgánica al nivel de partes por mil millones, sin embargo, sugiere que en Marte los compuestos orgánicos son activamente destruidos, probablemente por la fuerte radiación ultravioleta del Sol.

Los otros experimentos a bordo de los módulos de descenso Viking estaban dirigidos no ya a la búsqueda de materia





EL EXPERIMENTO DE INTERCAMBIO DE GASES probó el suelo marciano para ver si había en él microorganismos que absorbiesen gases atmosféricos y sustancias nutritivas y desprendiesen productos secundarios gaseosos. El experimento se llevó a cabo en dos fases. En la primera, se inyectó en la cámara un pequeño volumen de una solución nutritiva complicada *(color oscuro)* **de suerte que humedeciese la cámara sin mojar el suelo** *(arriba).* **Los gases desprendidos** *(gris claro)* **se transportaron a un cromatógrafo de gases mediante una corriente de helio** *(color claro);* **allí se analizaron en busca de compuestos orgánicos y se compararon con resultados de análisis de referencia efectuados como patrón. En la segunda fase de ese mismo experimento, se vertió en la cámara un gran volumen de solución nutritiva para mojar el suelo** *(parte inferior de la ilustración).*

orgánica en el suelo, sino a la de organismos vivientes. En la Tierra, los microorganismos (bacterias, levaduras y mohos) son los que pueden vivir bajo las condiciones más duras. Existen pocos lugares en la Tierra donde no haya formas de vida microbiana; son los últimos supervivientes en zonas de extremada temperatura y aridez. Las razones de su resistencia son interesantes, pero no hace falta detenernos aquí en ellas. Baste decir que, si hay vida en Marte, la probabilidad de detectarla se haría máxima dirigiendo la búsqueda a microorganismos en el suelo marciano.

Cada módulo de descenso Viking llevaba tres instrumentos ideados para detectar las actividades metabólicas de microorganismos en el suelo. Primero, el experimento de intercambio de gases, ideado para detectar cambios en la composición de la atmósfera producidos por metabolismo microbiano. Segundo, el experimento de desprendimiento marcado, proyectado para detectar la descomposición de compuestos orgánicos por los microbios del suelo cuando se les alimentaba con un nutriente. Tercero, el experimento de desprendimiento pirolítico, cuyo fin era detectar la síntesis de materia orgánica en el suelo marciano a partir de los gases de la atmósfera, bien por procesos fotosintéticos o no fotosintéticos. Los tres experimentos analizaron porciones de cada muestra de suelo marciano.

Todos los experimentos detectaron alteraciones químicas de una u otra clase

en el suelo. Todos los experimentos se han completado ya, y algunas de las alteraciones observadas sugieren procesos biológicos. Se ha examinado detenidamente, dentro y fuera del grupo de investigadores del proyecto Viking, cuál era la mejor manera de interpretar los hallazgos. ¿Se deben las alteraciones a respuestas biológicas o son sólo reacciones químicas que nos gustaría creer biológicas? Por otra parte, si la vida es una forma de química, ¿cómo se puede discernir una de la otra?

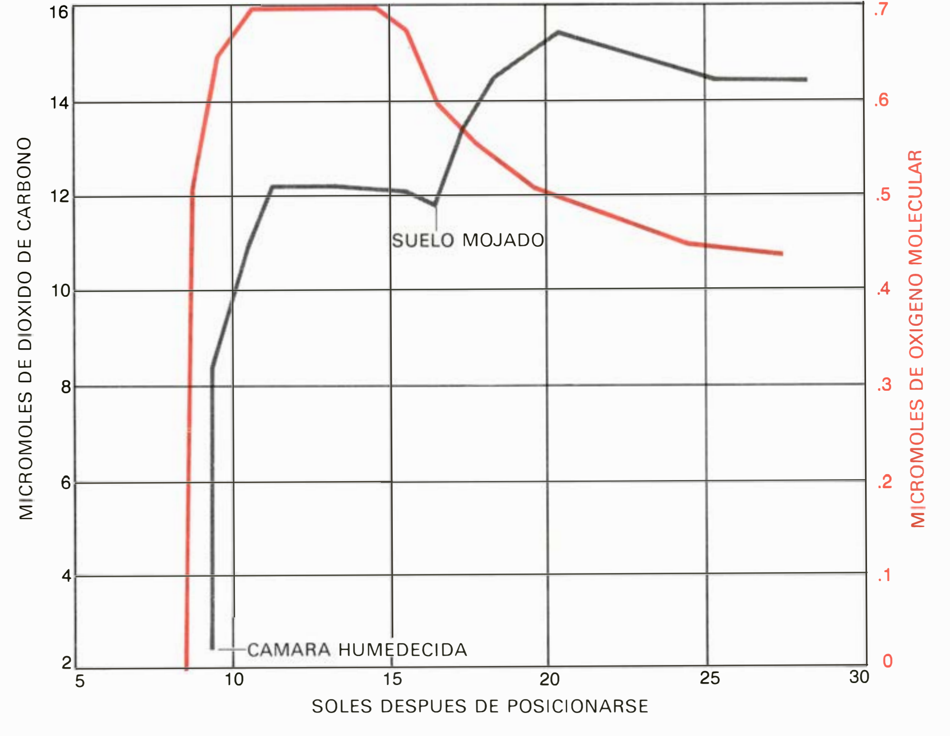
Un modo de decidir si un proceso es o no biológico consiste en probar su sensibilidad al calor. Las estructuras vivientes son frágiles y altamente organizadas, y se destruyen a temperaturas que apenas si afectan a muchas reacciones químicas. Por tanto, un proceso insensible al calor será con gran probabilidad una reacción química no viviente; ahora bien, un proceso sensible al calor podría ser o no de origen biológico.

La decisión respecto a si un proceso sensible al calor es o no biológico debe basarse en pruebas adicionales. A fin de cuentas, el juicio se basa en la navaja de Occam: el tradicional principio de que la hipótesis más probablemente correcta es aquella que da cuenta del máximo número de observaciones con el mínimo número de hipótesis.

El experimento de intercambio de gases y el de desprendimiento marcado eran de orientación francamente terrestre. En ambos experimentos, un medio nutritivo hecho de una solución acuosa de compuestos orgánicos se mezcló con una muestra de suelo marciano. Como el agua líquida no puede existir en Marte, el experimento no se podía llevar a cabo bajo condiciones marcianas; las cámaras de experimentación tuvieron que calentarse para evitar que el agua se congelara y acondicionarse a presión para evitar que hirviera. Ambos experimentos se basaban en la propiedad universal de los organismos terrestres de producir gas a medida que metabolizaban alimentos. Si una muestra de suelo terrestre se humedece con una solución nutritiva, los microorganismos del suelo absorben los nutrientes y los convierten parcialmente en más microorganismos (es decir, la población de microorganismos crece) y parcialmente en varios productos secundarios, incluyendo gases. Entre los gases desprendidos en el metabolismo microbiano están el dióxido de carbono, metano, nitrógeno, hidrógeno y sulfuro de hidrógeno. En la Tierra, los gases desprendidos por una especie de organismos son, tarde o temprano, consumidos por otras. De esa manera, los elementos ligeros en la superficie terrestre describen continuamente un ciclo a través de la biosfera y la atmósfera.

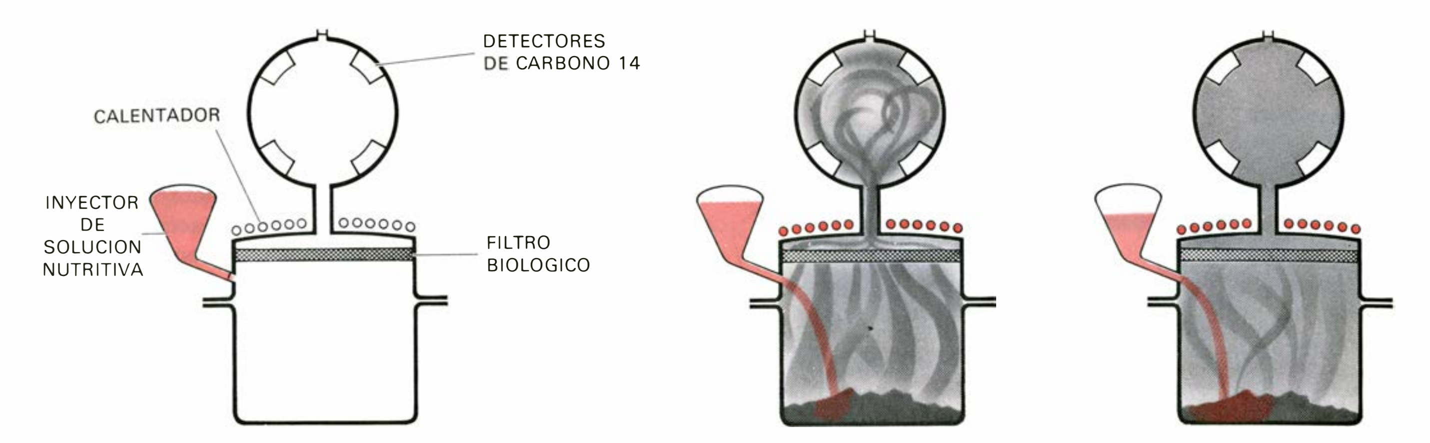
En el experimento de intercambio de gases, una solución nutritiva complicada se añadía a una muestra de suelo marciano en una cámara cerrada y los gases resultantes se analizaban periódicamente mediante un cromatógrafo de gases. El experimento se llevó a cabo en dos fases. En la primera se introducía en la cámara que contenía la muestra de suelo un pequeño volumen de la solución nutritiva, de manera que humedeciese la cámara sin mojar la muestra, y los gases resultantes se analizaban varias veces. En la segunda fase se echaba en la cámara un gran volumen de solución nutritiva, saturando la muestra de suelo. Con ésta ahora en contacto directo con el medio, comenzaba la parte principal del experimento. El suelo se incubaba durante casi siete meses, de modo que cualesquiera microorganismos que en él existieran tenían tiempo suficiente de dar señales de su presencia produciendo o consumiendo gases. Durante el tiempo de incubación se analizaba periódicamente la atmósfera en la cámara.

Los resultados de la primera fase del experimento fueron al mismo tiempo sencillos y sorprendentes. Inmediatamente después que la muestra de suelo se hubiera humedecido, hubo un rápido desprendimiento de dióxido de carbono y oxígeno. El desprendimiento de gases cesó al poco de haber comenzado, pero no antes de que la presión en la cámara hubiera aumentado apreciablemente. En el punto Chryse, en un período de un poco más de un sol, la cantidad de dióxido de carbono en la cámara de incubación del módulo *Viking I* se multiplicó por un factor cinco, y la cantidad de oxígeno por un factor 200. En el punto de Utopia, los aumentos fueron menores pero todavía considerables.

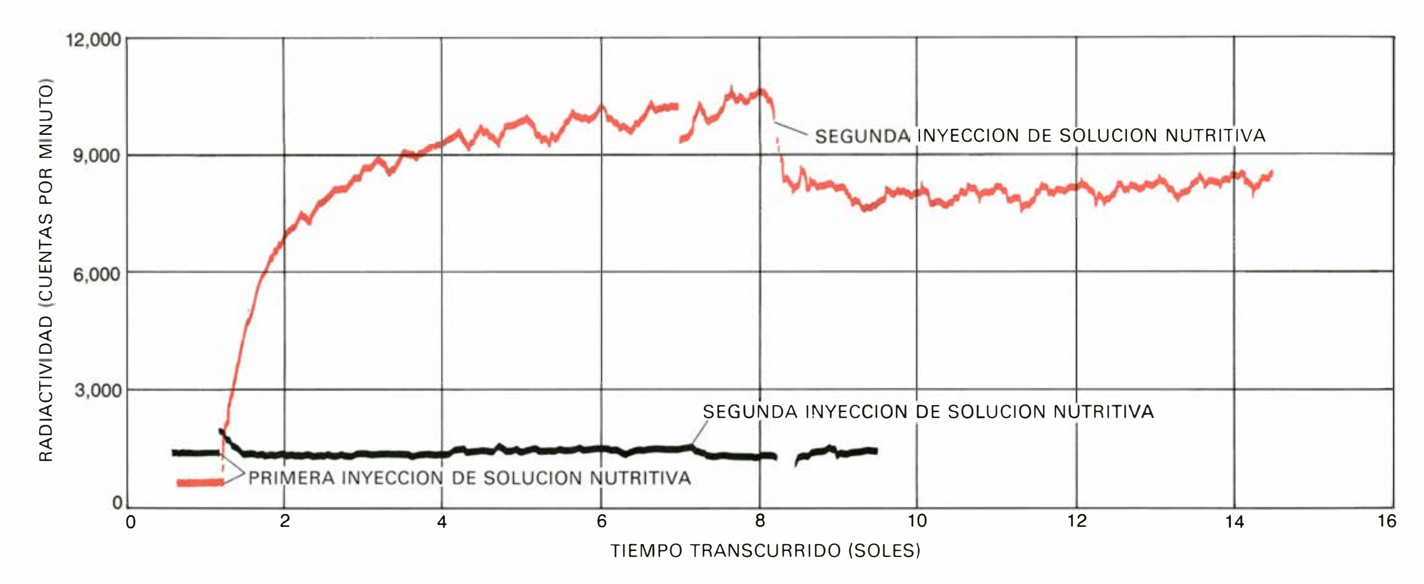


RESULTADOS DEL EXPERIMENTO de intercambio de gases, de acuerdo con datos de Vance I. Oyama, del Centro de Investigación Ames de la Administración Nacional de Aeronáutica y del Espacio (NASA). Ponen en evidencia que, en la primera fase húmeda del experimento, pasó a la cámara una gran cantidad de dióxido de carbono *(negro)* y oxígeno molecular *(color)*. En la segunda fase mojada, la cantidad de dióxido de carbono continuó aumentando a ritmo más reducido y descendió luego. La cantidad de oxígeno, sin embargo, decreció rápidamente. Se cree que los gases se desprendieron como consecuencia de procesos físicos y químicos, no de procesos biológicos. Un micromol es una millonésima de mol, siendo un mol una cantidad de una sustancia cuyo peso en gramos es igual a su peso molecular. La curva del oxígeno está trasladada hacia la izquierda en un sol de tal manera que las curvas no se superpongan. Se denomina sol a un día marciano.

La rapidez y brevedad de las respuestas registradas por ambos módulos de descenso sugería ciertamente que el proceso observado era una reacción química y no una reacción biológica. La aparición de dióxido de carbono se explica fácilmente. Era de esperar que se hubiera adsorbido gas dióxido de carbono en la superficie del seco suelo marciano; si el suelo se hubiera puesto en contacto con una atmósfera muy húmeda, el gas resultaría desalojado por el vapor de agua. La aparición del oxígeno es más complicada. La producción de tanto oxígeno parece requerir una reacción quí-



EL EXPERIMENTO DE DESPRENDIMIENTO MARCADO en los módulos de descenso Viking investigó el suelo marciano en busca de microorganismos que pudieran metabolizar compuestos orgánicos, o del carbono, sencillos. El medio nutritivo constaba de varios compuestos orgánicos ampliamente abundantes en el universo. Los compuestos se marcaron con carbono radiactivo. Si hubiese microorganismos en el suelo marciano, podrían consumir la solución nutritiva marcada y desprender gases radiactivos (particularmente dióxido de carbono), que detectarían los contadores de carbono 14. Antes de ensayar el suelo, se midió el nivel de radiación de fondo *(izquierda).* El suelo se introdujo en la cámara de prueba, se inyectó una pequeña cantidad de medio *(centro)* y se incubó hasta 11 soles. La cantidad de medio nutritivo en esta primera inyección se había escogido para mojar sólo una pequeña parte de la muestra, y humedecer la totalidad de la cámara. Una inyección subsiguiente del medio nutritivo *(derecha),* controlada por radio desde la Tierra, puso el medio en contacto con el suelo que ya se había mojado y con otro suelo que se había humedecido pero no mojado. Si el experimento de desprendimiento marcado se desarrollaba en Marte como se había planeado, sus resultados servirían de comprobación de los resultados del experimento de intercambio de gases, antes ilustrado.

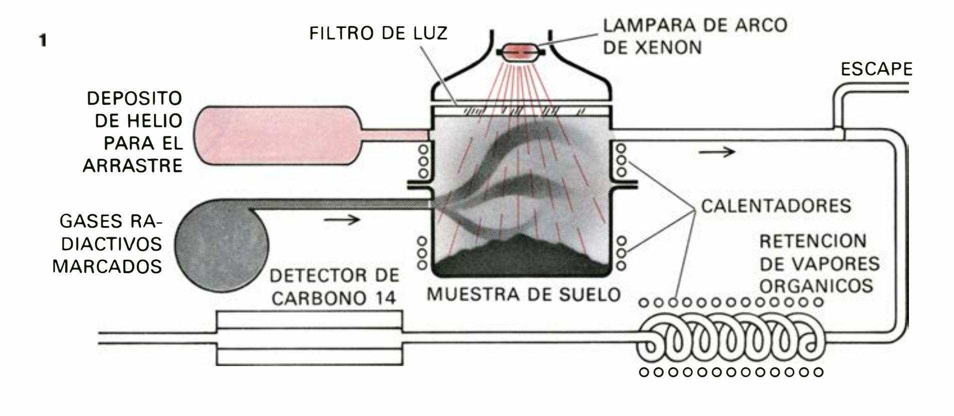


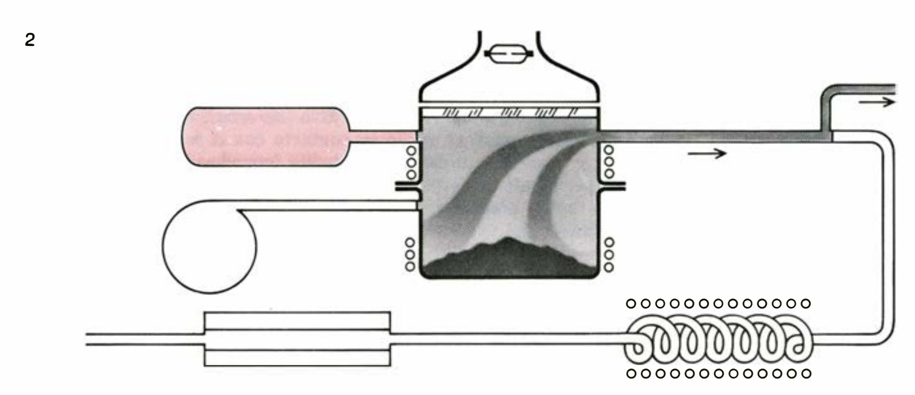
LOS RESULTADOS del experimento de desprendimiento marcado para la primera muestra de suelo analizado en el lugar de Chryse *(curva en color)* estaban en completo acuerdo con los resultados del experimento de intercambio de gases. Inmediatamente después de la primera inyección de medio nutritivo, hubo un fuerte desprendimiento de gases en la cámara. La radiactividad se midió con intervalos de 16 minutos durante todo el experimento, excepto las dos primeras horas tras la primera inyección, en que las mediciones se hicieron cada cuatro minutos. Después de la segunda inyección, la cantidad de gas en la cámara descendió, permaneciendo después a un nivel casi constante hasta el final del experimento. Para probar la sensibilidad de la reacción al calor, se esterilizó una segunda muestra de suelo a una temperatura de 160 grados Celsius durante tres horas, y después se repitió el experimento *(curva en negro).* La reacción no se produjo. Aunque tal comportamiento concuerda con el de un proceso biológico, es más probable que, de nuevo, el experimento detectase solamente una reacción química.

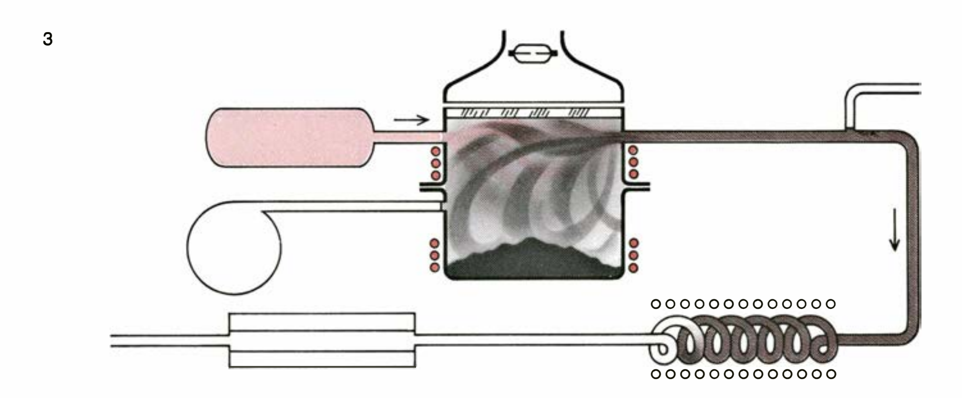
mica que lo produzca y no sólo una liberación física de gas preexistente. Probablemente, el oxígeno se desprendió cuando el vapor de agua descompuso un compuesto rico en oxígeno, tal como un peróxido. Se sabe que éstos se descomponen si se ponen en contacto con agua en presencia de compuestos de hierro, y, según el espectrómetro de radiofluorescencia a bordo de cada módulo de descenso Viking, el suelo marciano contiene un 13 por ciento de hierro.

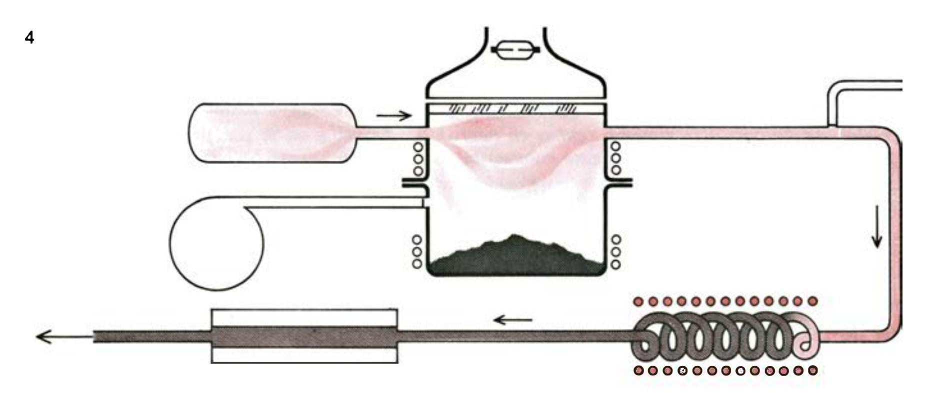
En ambos puntos de descenso, la segunda fase del experimento de intercambio de gases supuso una decepción. Cuando la muestra de suelo se saturó con medio nutritivo y se incubó, continuó desprendiéndose anhídrido carbónico. La producción de éste disminuyó gradualmente, y el oxígeno desapareció poco a poco. El lento aumento en la cantidad de anhídrido carbónico era, probablemente, una continuación de la reacción en la fase húmeda del experimento. La desaparición del oxígeno puede también explicarse sin dificultad: uno de los ingredientes del medio nutritivo era ácido ascórbico, que se combina fácilmente con oxígeno. Por tanto, resultó claro que después de varios meses, todo lo interesante había ocurrido en la fase húmeda del experimento, antes de que el suelo entrara en contacto con el nutriente. Lo que el experimento de intercambio de gases detectó no era metabolismo, sino la interacción química del material de la superficie marciana con vapor de agua a una presión que no se había alcanzado en Marte en muchos millones de años.

El experimento de desprendimiento marcado difería del de intercambio de gases en varios respectos. El medio nutritivo empleado era simple; contenía unos cuantos compuestos cósmicamente









EXPERIMENTO DE DESPRENDIMIENTO PIROLITICO, que sometió a ensayo el suelo marciano *(gris oscuro)* para comprobar si había en él microorganismos que creasen compuestos orgánicos a partir de los gases atmosféricos, bien por un proceso fotosintético o bien por un proceso no fotosintético. Se encerró herméticamente una muestra de suelo en la cámara junto con algo de atmósfera marciana y una pequeña cantidad de monóxido y dióxido de carbono radiactivos *(gris claro)*. Una lámpara de arco de xenón irradiaba el suelo con luz solar marciana simulada (*1*). A los cinco días, se apagó la lámpara y se retiró la atmósfera de la cámara (*2*). El suelo se calentó a una temperatura suficientemente alta para pirolizar (descomponer) en fragmentos pequeños y volátiles cualquier compuesto orgánico radiactivo producido. Los fragmentos *(gris claro)* se arrastraron fuera de la cámara (*3*) mediante una corriente de helio *(color claro)* hasta una columna ideada para retener moléculas orgánicas, dejando pasar dióxido y monóxido de carbono. Las moléculas orgánicas radiactivas retenidas en la columna fueron liberadas elevando la temperatura de la columna; las moléculas se oxidaron para formar dióxido de carbono, que pasó al contador de radiación *(4)*.

abundantes como el ácido fórmico (HCOOH) y el aminoácido glicina (NH_2CH_2COOH). Todos los compuestos estaban marcados con átomos del isótopo radiactivo ^{14}C. El instrumento de desprendimiento marcado se ideó para detectar gases radiactivos, principalmente dióxido de carbono, desprendidos cuando el medio nutritivo se añadía a una muestra de suelo. El número de desintegraciones radiactivas en gases puede contarse de manera muy eficiente, de modo que el experimento de desprendimiento marcado es más rápido y sensible que el de intercambio de gases para detectar actividad microbiana en el suelo terrestre. La secuencia de operaciones del experimento de desprendimiento marcado no incluía una fase húmeda como tal, pero pretendía lograr el mismo fin inyectando un volumen de medio nutritivo que fuese insuficiente para mojar la totalidad de la muestra de suelo, pero suficiente para humedecer la cámara. Si el experimento se desarrollase en Marte en la forma prevista, inyecciones subsiguientes del medio, controladas por órdenes enviadas desde la Tierra, pondrían el medio en contacto con porciones de suelo que se habían mojado previamente y con otras que se habían humedecido, pero no mojado.

Como en el experimento de intercambio de gases, inmediatamente después de añadir medio nutritivo al suelo en el experimento de desprendimiento marcado se produjo fuerte desprendimiento de gas en la cámara. Este desprendimiento disminuyó gradualmente poco después del primer sol. El gas (dióxido de carbono indudablemente) era radiactivo, lo que demostraba que se había formado de los compuestos radiactivos del medio y no de compuestos del suelo marciano. Los gases no radiactivos, que también debieron formarse cuando el medio acuoso entró en contacto con el suelo, no eran detectables en el experimento.

La producción de dióxido de carbono radiactivo en el experimento de desprendimiento marcado se comprende a la luz del resultado del experimento de intercambio de gases, y sugiere que el suelo marciano contiene peróxidos. El ácido fórmico, uno de los compuestos del medio en el desprendimiento marcado, se oxida con especial facilidad: si una molécula de ácido fórmico (HCOOH) reacciona con otra de peróxido de hidrógeno (H_2O_2), se formará una molécula de dióxido de carbono (CO_2) y dos de agua ($2H_2O$). La cantidad de dióxido de carbono radiactivo liberado en el experimento de desprendimiento marcado fue sólo un poco menor de lo que se espera-

ría si todo el ácido fórmico en el medio se hubiese oxidado en esta forma.

Si la fuente del oxígeno desprendido en la fase húmeda del experimento de intercambio de gases fuesen en realidad peróxidos descompuestos en el suelo por el vapor de agua, entonces en el experimento de desprendimiento marcado todos los peróxidos deberían también haber sido descompuestos por la primera inyección de nutriente. Por tanto, la siguiente inyección no debería haber producido desprendimiento adicional de gas radiactivo, pese a que parte de la muestra presumiblemente no había sido mojada aún por el medio. Eso fue lo que ocurrió. Cuando se inyectó en la cámara un segundo volumen de medio, la cantidad de gas en la cámara no aumentó; en realidad, disminuyó. La disminución encuentra su explicación en que el dióxido de carbono es muy soluble en agua; cuando se añadió a la cámara nuevo medio nutritivo, absorbió parte del dióxido de carbono presente en la zona no ocupada por la muestra.

Este resultado se obtuvo con todas las muestras sometidas a prueba mediante el experimento de desprendimiento marcado, en ambos puntos de descenso de los Viking. En ese aspecto, los resultados del experimento de desprendimiento marcado no fueron paralelos a los del experimento de intercambio de gases. En ambos lugares, uno y otro experimento pusieron a prueba suelo recogido de la superficie; en el caso de Utopia, los experimentos también se efectuaron con suelo recogido debajo de una roca. Aunque el experimento de desprendimiento marcado no encontró diferencia notable en la cantidad de gas desprendido por cualquiera de las muestras, el experimento de intercambio de gases registró, en las muestras procedentes de Utopia, alrededor de tres cuartas partes de la cantidad de dióxido de carbono recogida en el lugar Chryse, y todavía menos dióxido de carbono en la muestra recogida debajo de la roca.

El experimento de intercambio de gases también registró menos oxígeno de las muestras de la región de Utopia, pero la interferencia del ácido ascórbico presente en el medio nutritivo de ese experimento hace difícil cuantificar la diferencia. En todo caso, el experimento de intercambio de gases detectó considerablemente mayores cantidades de gas que el experimento de desprendimiento marcado para porciones de la misma muestra. Esos resultados, sin embargo, no contradicen la tesis de que la producción de oxígeno detectada por el experimento de intercambio de gases y la producción de dióxido de carbono radiactivo detectada por el experimento de desprendimiento marcado son simplemente mediciones diferentes de la misma química de la superficie. El experimento de intercambio de gases mide la cantidad total de oxidante en la superficie; el de desprendimiento marcado mide sólo una fracción de ella.

El experimento de desprendimiento marcado investigó también la estabilidad de la reacción respecto al calor. Cuando el suelo se precalentó a 160 grados centígrados durante tres horas antes de la incubación, la reacción quedó eliminada. Cuando se calentó a 46 grados durante el mismo tiempo, la magnitud de la reacción se redujo a aproximadamente la mitad. Estos resultados se han considerado por algunos como prueba en favor de la hipótesis de que la reacción es biológica. Desde luego, los resultados son coherentes con tal hipótesis, pero también lo son con una oxidación química en que el agente oxidante se destruya o evapore a una temperatura relativamente baja. A esos mismos resultados se podría haber llegado, probablemente, con una gran variedad de peróxidos, tanto inorgánicos como orgánicos.

El tercer experimento microbiológico, el de desprendimiento pirolítico, difería del de intercambio de gases y del de desprendimiento marcado en dos puntos. Primero, pretendía medir la síntesis de materia orgánica a partir de gases atmosféricos, más bien que la descomposición. Segundo, se ideó para operar bajo las condiciones de temperatura, presión y composición atmosféricas que realmente se dan en Marte, ya que ésas son las condiciones bajo las que debe existir cualquier forma de vida marciana. En la práctica, las condiciones de la cámara constituían una aproximación razonablemente buena de las condiciones marcianas, excepto por lo que respecta a la temperatura, que se mantuvo más alta que la exterior por existir fuentes caloríficas en el interior de la nave.

Se encerró herméticamente, en una cámara, una muestra de suelo marciano junto con algo de atmósfera marciana. Una ventana de cuarzo en la cámara permitía el paso de luz solar marciana, simulada por una lámpara de arco de xenón. Dentro de ese pequeño microcosmos marciano se introdujeron pequeñas cantidades de dióxido de carbono y monóxido de carbono, ambos radiactivos. Los dos gases se hallan en la atmósfera marciana, pero no en forma radiactiva. A los cinco días se apagó la lámpara, se extrajo la atmósfera de la cámara y se analizó la muestra de

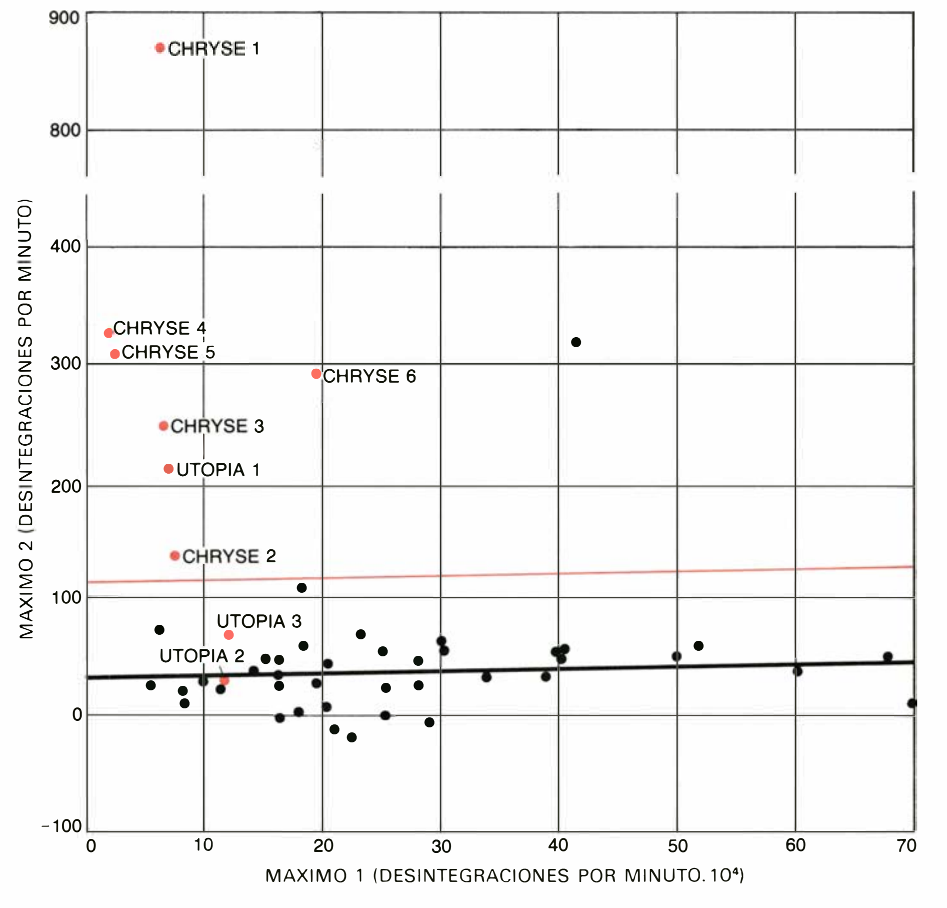
suelo en busca de presencia de materia orgánica radiactiva.

Se calentó primero la muestra en el horno de pirolisis, a una temperatura suficientemente alta para provocar la rotura de cualquier molécula de compuesto orgánico en pequeños fragmentos volátiles. Se arrastraron los fragmentos fuera de la cámara mediante una corriente de helio y se hicieron pasar a través de una columna proyectada para atrapar moléculas orgánicas, aunque permitiendo el paso de monóxido y dióxido de carbono. Las moléculas orgánicas radiactivas se traspasaron así desde el suelo a la columna; al mismo tiempo, se separaron de cualquier gas remanente de la atmósfera de incubación. Las moléculas orgánicas se desprendieron después de la columna elevando la temperatura de ésta. Simultáneamente, las moléculas orgánicas radiactivas se descompusieron en dióxido de carbono radiactivo, en presencia de óxido de cobre en la columna. La corriente de helio arrastró entonces el dióxido de carbono a un contador de radiación. Si se hubieran sintetizado compuestos orgánicos en el suelo, se detectarían en forma de dióxido de carbono radiactivo; si no se hubieran sintetizado compuestos orgánicos, no se habría formado dióxido de carbono radiactivo.

Sorprendentemente, siete de los nueve experimentos de desprendimiento pirolítico realizados en Marte dieron resultados positivos. Los dos resultados negativos se obtuvieron en el lugar de Utopia, pero una tercera muestra ensayada en Utopia fue positiva. Esta tercera muestra se incubó en la oscuridad, lo que implica que la luz puede no ser necesaria para la reacción. La cantidad de carbono fijado en el suelo por el experimento fue pequeña: suficiente para suministrar materia orgánica para entre 100 y 1000 células bacterianas. Se trata, efectivamente, de una cantidad mínima, hasta el extremo de que podría haberse escapado a la detección en el experimento de análisis orgánico. No obstante, esa cantidad resulta significativa; sorprendió el que, en un ambiente tan fuertemente oxidante, pudiera fijarse en el suelo incluso una pequeña cantidad de materia orgánica.

Más significativo es el hecho de que el experimento de desprendimiento pirolítico se hubiera proyectado rigurosamente para eliminar fuentes no biológicas de compuestos orgánicos. Durante el desarrollo del experimento se había encontrado que, en presencia de radiación ultravioleta de muy corta longitud de onda, el monóxido de carbono se combinaba espontáneamente con vapor de agua para formar moléculas orgánicas sobre las superficies de vidrio, cuarzo y suelo en la cámara experimental. Para evitar esas reacciones y la confusión que habrían causado, el ultravioleta de longitud de onda corta se filtró de la radiación a la que se permitía penetrar en la cámara de incubación. Recibir resultados positivos del suelo de Marte, a pesar de esas precauciones, fue algo asombroso.

Sin embargo, parece que los hallazgos del experimento de desprendimiento pirolítico deben también interpretarse en términos no biológicos. La razón es que la reacción detectada resultó menos sensible al calor de lo que debería esperarse en un proceso biológico. En dos de los nueve ensayos de desprendimiento pirolítico realizados en Marte, la muestra de suelo se calentó antes de que se inyectasen los gases radiactivos y comenzase la incubación. En un caso la muestra se mantuvo a 175 grados centígrados durante tres horas y, en el otro, se mantuvo a 90 grados durante casi dos horas. La temperatura más elevada redujo la reacción en casi un 90 por ciento, pero no la eliminó. La temperatura más baja no aportó ninguna novedad. Cuando se recuerda que la temperatura de la superficie de Marte en los dos lugares de toma no sube por encima de cero grados Celsius en ningún caso, y que la temperatura debajo de la superficie es todavía más baja, se hace difícil reconciliar los resultados con una fuente biológica. Todos los organismos vivientes en el suelo marciano deberían haber muerto a esas temperaturas.



LOS RESULTADOS del experimento de desprendimiento pirolítico se representan aquí para todas las muestras ensayadas en Marte *(puntos en color)*. El eje rotulado "Máximo 1" muestra la cantidad de radiactividad en forma de dióxido y monóxido de carbono que pasó a través de la columna durante la pirolisis de cada muestra. El eje rotulado "Máximo 2" muestra la cantidad de radiactividad que permaneció ligada a la columna en cada caso, representando materia orgánica recién sintetizada. Cada punto en color está rotulado y numerado según el lugar en el que fue ensayado y con el experimento que representa. (Por ejemplo, "Chryse 1" significa el resultado del primer experimento en Chryse, "Utopia 1" representa el resultado del primer experimento Utopia, y así sucesivamente.) Los puntos en negro son los datos obtenidos de ensayos de muestras de suelo esterilizadas en un duplicado del instrumento Viking de desprendimiento pirolítico en la Tierra. La línea negra trazada por esos puntos representa la recta de mejor adaptación a ellos. La línea en color por encima de la negra es una línea divisora de representatividad estadística; cualquier punto situado por encima de la línea en color constituye un resultado positivo. Se cree que el punto negro aislado encima de la línea en color se debe a un error técnico al efectuar ese ensayo particular. Siete de los nueve experimentos de desprendimiento pirolítico realizados en Marte, sin embargo, produjeron resultados positivos.

Por otra parte, no es fácil configurar una explicación no biológica de los resultados positivos. Se están llevando a cabo investigaciones sobre el problema en laboratorios terrestres con suelos marcianos sintéticos, construidos a base de los datos procedentes de los análisis inorgánicos realizados por los módulos Viking de entrada. La solución del rompecabezas explicará también, probablemente, por qué el experimento de análisis orgánico no detectó materia orgánica en el suelo de Marte. Hasta que el misterio de los resultados del experimento de desprendimiento pirolítico se resuelva, la explicación biológica continuará siendo una posibilidad remota.

Aun cuando quedan algunas ambigüedades, no son muchas las incógnitas por despejar en torno al significado de las observaciones de los módulos de descenso Viking: por lo menos las zonas de Marte que han examinado los dos vehículos espaciales no son hábitats de vida. Posiblemente la misma conclusión se aplica a la totalidad del planeta, pero eso es un problema complicado que todavía no puede abordarse. El resultado más sorprendente del experimento en busca de la vida es la extraordinaria reactividad química del suelo marciano: su capacidad oxidante, su falta de materia orgánica hasta el nivel de partes por mil millones y su capacidad para fijar carbono atmosférico (presumiblemente en moléculas orgánicas) a un nivel todavía más bajo. Según parece, Marte tiene una superficie fotoquímicamente activada que, debido a la baja temperatura y la ausencia de agua, se mantiene en un estado lejano del equilibrio químico.

Estas conclusiones deducidas de los resultados de los experimentos de búsqueda de vida en Marte son innegablemente decepcionantes. Cuando menos, el descubrimiento de vida habría sido mucho más interesante. Hay todavía personas que, poco dispuestas a aceptar la idea de un Marte sin vida, sostendrán que la interpretación expuesta no queda demostrada. Tienen razón. Es imposible demostrar que cualquiera de las reacciones detectadas por los instrumentos de los Viking no tienen origen biológico. Es igualmente imposible demostrar, partiendo de alguno de los resultados de los experimentos Viking, que las rocas observadas en los lugares de descenso no eran organismos vivientes que casualmente parecían rocas. En cuanto se abandona la navaja de Occam, el campo queda abierto a cualquier fantasía. Siglos de experiencia humana nos advierten, sin embargo, de que tal enfoque no es el camino para descubrir la verdad.

Riego gota a gota

En este sistema de tubos de plástico fijados a la superficie del suelo se suministra agua a las plantas gota a gota. El sistema reduce la tensión de la planta, mantiene la humedad y puede operar con agua salina

Kobe Shoji

Hace unos 40 años, el ingeniero israelita Symcha Blass observó que el crecimiento de un gran árbol cercano a un grifo que goteaba era más vigoroso que el de los árboles de la misma zona no alcanzados por el agua del grifo. Blass sabía que los métodos de riego clásicos desperdiciaban gran cantidad del agua aplicada a los cultivos; por eso, el ejemplo del grifo que goteaba fue el chispazo que le hizo pensar en un sistema de riego que suministrara agua en cantidades pequeñas, literalmente gota a gota. Con el tiempo, llegaría a diseñar y patentar un sistema de baja presión que vertía pequeñas cantidades de agua a las raíces de las plantas, a intervalos frecuentes. La técnica, tal como la desarrolló Blass, y fue posteriormente perfeccionada por él y distintos fabricantes, consiste en fijar un tubo de plástico pequeño a la superficie del campo, junto a las plantas, para que fluya agua lenta y continuamente desde agujeros o emisores especiales situados en los puntos apropiados a lo largo del tubo. Ese sistema, llamado actualmente riego gota a gota, o por goteo, ha alcanzado mucha aceptación. Se ha mostrado particularmente eficaz en zonas que son áridas y tienen costes de mano de obra elevados. Una ventaja insospechada es que el sistema funciona bien incluso con agua muy salina, como suele ser la de regiones áridas.

En muchas explotaciones agrícolas del mundo se sigue regando con el mismo método que se hacía 5000 años atrás, a saber: por inundación de los campos o por desviación del agua hacia la cosecha mediante surcos paralelos. Con estas formas de riego las plantas toman únicamente del 30 al 60 por ciento del agua que se les suministra. Además, el acondicionamiento del campo y la maniobrabilidad del sistema de riego exige considerable trabajo. Si el agua abunda y el coste de mano de obra no sube demasiado, el método de los surcos puede tener cierta ventaja económica no obstante su ineficacia. Sin embargo, en algunas zonas, como California y Hawaii, que se enfrentan, no sólo ante una escasez de agua, sino también a un encarecimiento creciente de los braceros, la energía y el agua, el riego gota a gota se ha adaptado rápida y satisfactoriamente. Así, California, que hace dos años fue el tercero por su sequía, posee en la actualidad el 15 por ciento del total mundial, con 162.000 hectáreas, del sistema de riego gota a gota.

La idea fundamental, base del riego gota a gota, puede buscarse en ciertos experimentos llevados a cabo en Alemania en 1860. Los agricultores dejaban abiertas las junturas de las tuberías de arcilla a unos 0,8 metros de profundidad, en un esfuerzo por combinar el riego y el drenaje según el nivel de agua ascendía o descendía durante el año. En 1930, los colonos australianos, que disponían de unos recursos hídricos limitados, proyectaron un sistema de riego para huertas de melocotones mediante tuberías de hierro galvanizado de cinco centímetros, en las que se habían practicado agujeros con un punzón. Los trabajadores de invernaderos en el Reino Unido comenzaron a ensayar un método similar en 1948 para el cultivo de tomates.

Las técnicas del riego gota a gota no entrarían en Estados Unidos hasta principios de la década de 1960, cuando varios cultivadores de viveros instalaron el sistema en los invernaderos. En 1968, se empezó a aplicar a huertas y cultivos en surcos en California; a partir de entonces, se ha ido adaptando a otros cultivos, verbigracia, tomates, viñas, fresas, maíz, piña tropical y caña de azúcar. (El riego gota a gota no es adecuado para las plantas que se cultivan muy juntas, como cereales y alfalfa, ya que la cantidad de tubería que requieren hace el sistema antieconómico.)

Cuando Blass concibió la idea de un sistema de riego gota a gota por los años en torno a 1930, no se disponía de los materiales necesarios para construir una estructura de presión baja a un precio razonable. Hubo que esperar al rápido desarrollo de la industria del plástico, después de la Segunda Guerra Mundial, para que se fabricaran materiales apropiados para construir tubos flexibles, económicos y resistentes a los agentes químicos. Los primeros sistemas de riego gota a gota consistían en tubos capilares de plástico (de un milímetro de diámetro) que se acoplaban a otros tubos mayores. El rozamiento del tubo restringía el aflujo de agua al suelo entre dos y cuatro litros por hora, para cada punto de descarga. El sistema se instaló inicialmente bajo tierra, pero a causa de las primitivas técnicas de filtración de aquella época y la frecuente obturación del sistema de distribución se trasladó a la superficie. Este cambio facilitó la comprobación de los tubos obturados y mantuvo la principal ventaja del sistema, a saber: la aplicación directa del agua a la zona radicular de la planta.

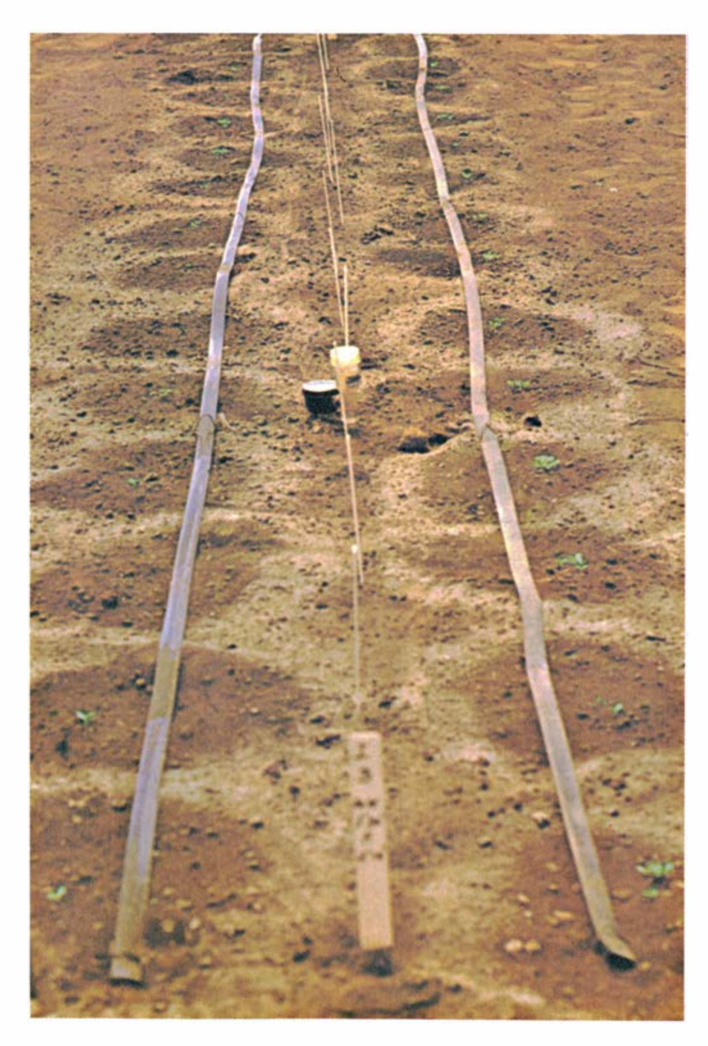
Entre los perfeccionamientos que hizo Blass de su sistema original se cuenta el diseño de un emisor circular que evitara el peligro de la obturación. Consistía en un tubo espiral encapsulado en un recubrimiento tenaz. El tubo servía para reducir la presión de descarga al prolongar la trayectoria del flujo del agua; ello permitía la descarga del agua por un agujero mayor.

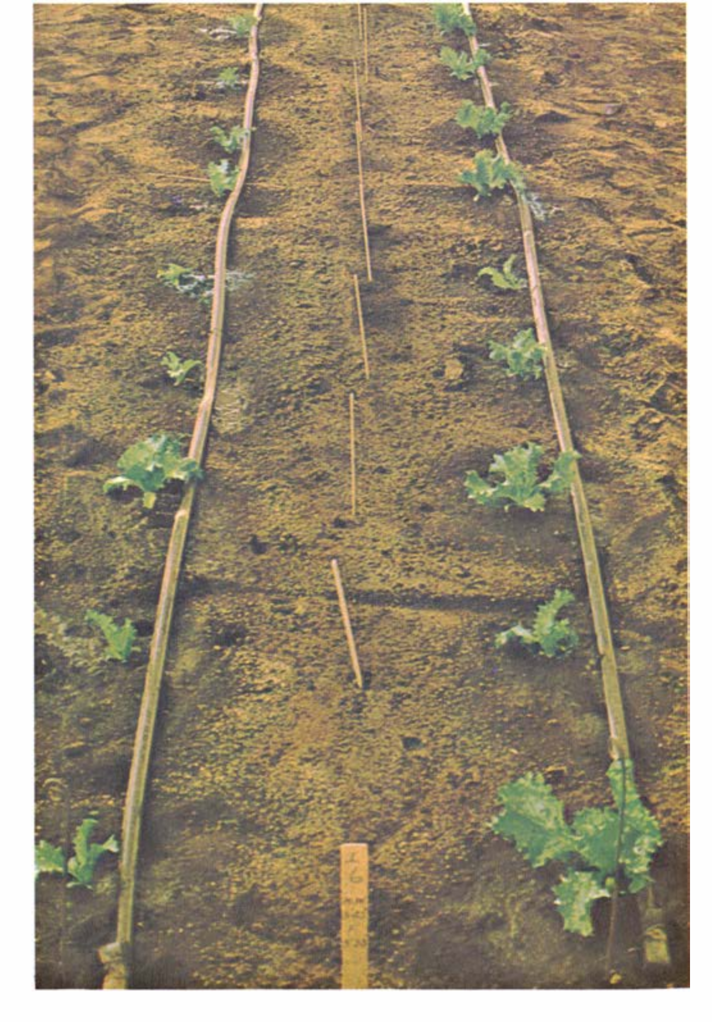
En 1960, los científicos israelitas que trabajaban en campos experimentales hicieron público su éxito espectacular obtenido al aplicar el sistema de Blass a zonas desérticas del Negev y de Arava. Las cosechas habían sido pobres con riego de surco y por aspersión. Lo cual se debía principalmente a que el agua era muy salina. El riego gota a gota puede utilizar, en términos generales, agua más salina de la que se aceptaría con otros métodos de riego, tema sobre el que volveremos.

Las condiciones para la agricultura en zonas desérticas eran inequívocamente adversas; así, además de la elevada salinidad del agua, estaba la alta temperatura, la humedad relativa baja y el suelo arenoso. A pesar de tales dificultades, la técnica del riego gota a gota consiguió una mejora sustancial del rendimiento de las cosechas. Por ejemplo, un campo experimental de Arava produjo cosechas anuales de casi 53,3 toneladas métricas de tomates de invierno por hectárea con el riego gota a gota, frente a las 35,8 toneladas métricas alcanzadas con el riego por aspersión. En otro ensayo, la producción de melones se incrementó en un 70 por ciento, y los cultivos sensibles a las sales, como los pepinos, registraron por vez primera una cosecha significativa. El riego gota a gota demostró otras ventajas adicionales: disminución del consumo de agua, reducción del coste del trabajo invertido y una mejor distribución del abono. Los fertilizantes solubles pueden distribuirse fácilmente con el agua a lo largo del sistema de riego; al depositar éste el agua sólo al lado de la planta, el abono se aplica allí donde es más necesario.

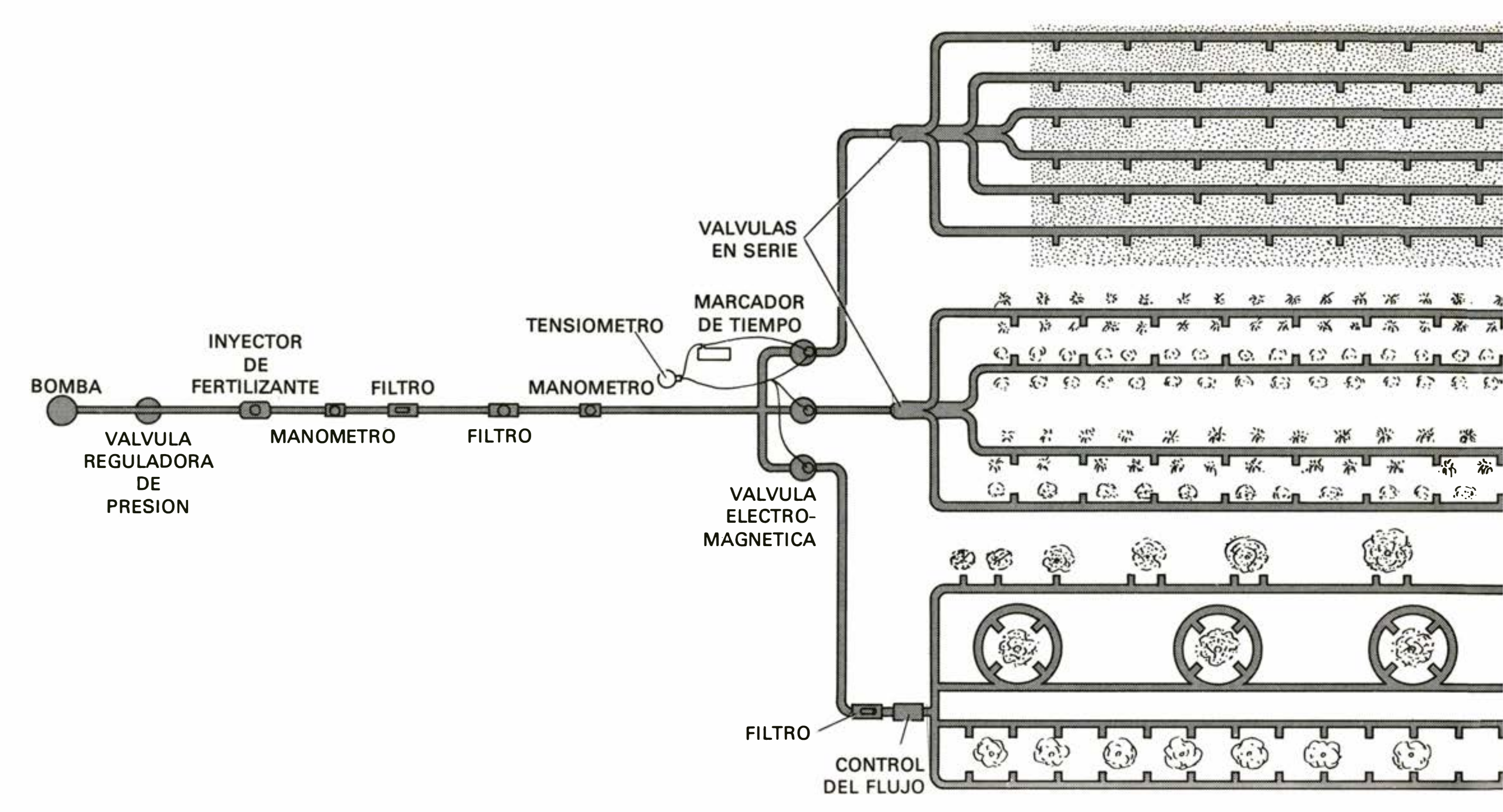
Desde la última década ha venido apareciendo toda una industria dedicada al diseño y manufactura de material para el riego gota a gota. Ello ha traído consigo un cambio inmediato en los bienes de equipo empleados. El riego gota a gota se aplica a una amplia variedad de suelos pertenecientes a diferentes topografías, de forma que el diseño de un sistema para mantener flujos uniformes comporta un reto para el ingeniero. Y aunque todo sistema gota a gota se basa en la hidrodinámica de las tuberías, sin embargo, cada instalación debe diseñarse específicamente para el cultivo al que aporta el flujo, de acuerdo con las condiciones bajo las cuales el cultivo se desarrolla.

Un típico sistema de riego gota a gota moderno consta de una red de tuberías de plástico y tubos graduables. Una tubería, de diámetro bastante grande, conduce el agua al borde del campo. Una serie de líneas de cañerías principales de diámetro menor la introducen dentro del campo, sin descargarla todavía en el cultivo. Otras cañerías, secundarias, de tamaño menor que las principales, la transportan a las líneas laterales, desde las cuales es aplicada a las plantas a través de emisores o agujeros. La línea de abastecimiento (de unos 30 centímetros de





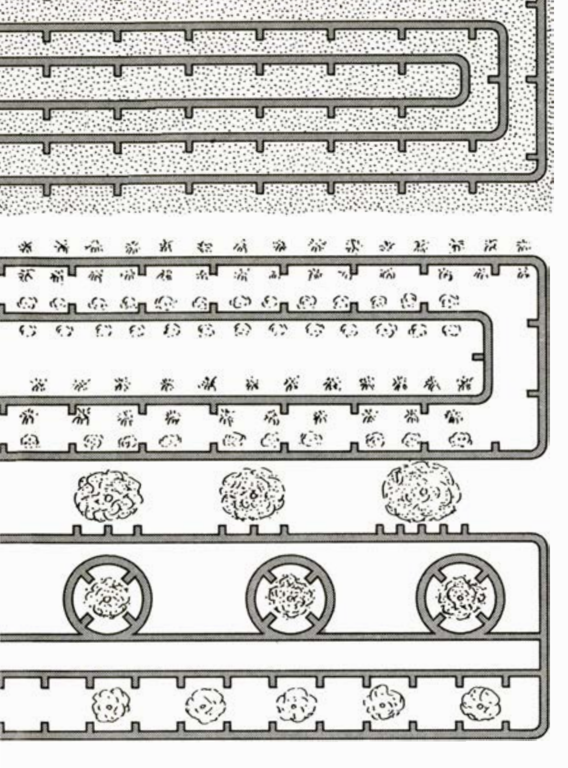
EL SISTEMA DE RIEGO GOTA A GOTA en un campo experimental de la Universidad de Hawaii. A la izquierda, tubos estrechos de plástico suministran agua a plántulas individuales mediante unos agujeros muy pequeños perforados a intervalos adecuados. Las manchas de humedad en el suelo indican la precisión con la que el agua se aplica. A la derecha, se observan las plantas de lechuga en una parcela cercana después de cultivarlas mediante el riego gota a gota durante cuatro semanas. En algunos sistemas más complejos se instalan emisores en las tuberías.



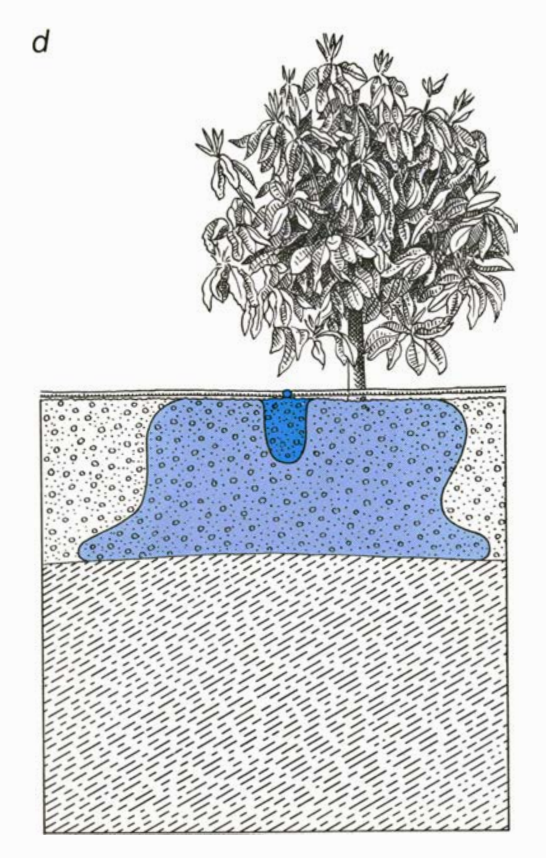
MODELOS DE TUBERIAS utilizados en el riego gota a gota en una configuración esquematizada. Los tubos y los emisores se disponen de forma diferente para distintos cultivos. En un campo, en donde los árboles son pequeños y están cerca unos de otros, las tuberías pueden extenderse, paralelamente a la hilera de árboles, con numerosos orificios o emisores. Con árboles grandes, las tuberías se enlazan alrededor de cada individuo, también con varios emisores. Para un cultivo en hilera, la configuración puede disponerse de acuerdo con una tubería por fila o una tubería cada dos filas. Cuando las plantas de un cultivo crecen muy juntas, como en el caso de la alfalfa y los cereales, el riego gota a gota puede no resultar el método idóneo, pues la cantidad de tuberías requeridas pueden hacerlo antieconómico. Si se utiliza el riego gota a

a b c

LA ZONA HUMEDECIDA alrededor de las raíces de un árbol o plantas regadas por el método gota a gota varía en su sección transversal según sea el suelo. Se coloca un emisor al lado del árbol; la zona más húmeda *(color obscuro)* se halla directamente debajo del emisor. En el suelo bien acondicionado, de textura fina *(a)*, predominan las fuerzas capilares, y el perfil de la zona húmeda *(color claro)* se parece a una media esfera. Si el suelo está pobremente acondicionado, los apelmazamientos y los terrones interfieren en la acción capilar, y el perfil



gota en estos sembrados, las tuberías y los emisores se espacian uniformemente. La ordenación de válvulas hace posible regar zonas distintas en serie. Los tensiómetros miden la humedad del suelo de manera que la distribución del agua pueda programarse de forma eficaz.



está más condicionado por la gravedad *(b)*. En suelos de textura arenosa *(c)* dominan las fuerzas gravitacionales. La capa dura *(d)* frena el descenso y favorece el movimiento lateral.

diámetro) y las líneas principales (15 centímetros) son tuberías de plástico rígido; se suelen instalar por debajo de la superficie. Si se incluyen cañerías secundarias, probablemente su diámetro varía entre 7,6 y 15,2 centímetros. En algunos sistemas, las líneas laterales están unidas directamente a la línea de la cañería principal.

Las líneas laterales, flexibles, miden de 12 a 32 milímetros de diámetro. Están instaladas al lado del plantel, en superficie o ligeramente soterradas. Los emisores acoplados a la línea lateral (agujeros espaciados a lo largo de ella) ceden agua a la planta a un ritmo de uno a seis litros por hora. En una huerta o viñedo típicos hay de uno a seis puntos de descarga de agua por planta; depende del tamaño de la misma y del intervalo entre una y otra. Los puntos se disponen en dirección paralela a la fila de las plantas o bien en círculo alrededor de cada planta. Para un cultivo en fila, las líneas laterales están constituidas por mangueras porosas o por tubos con orificios del tamaño de un alfiler, y se instalan paralelas a la dirección del cultivo.

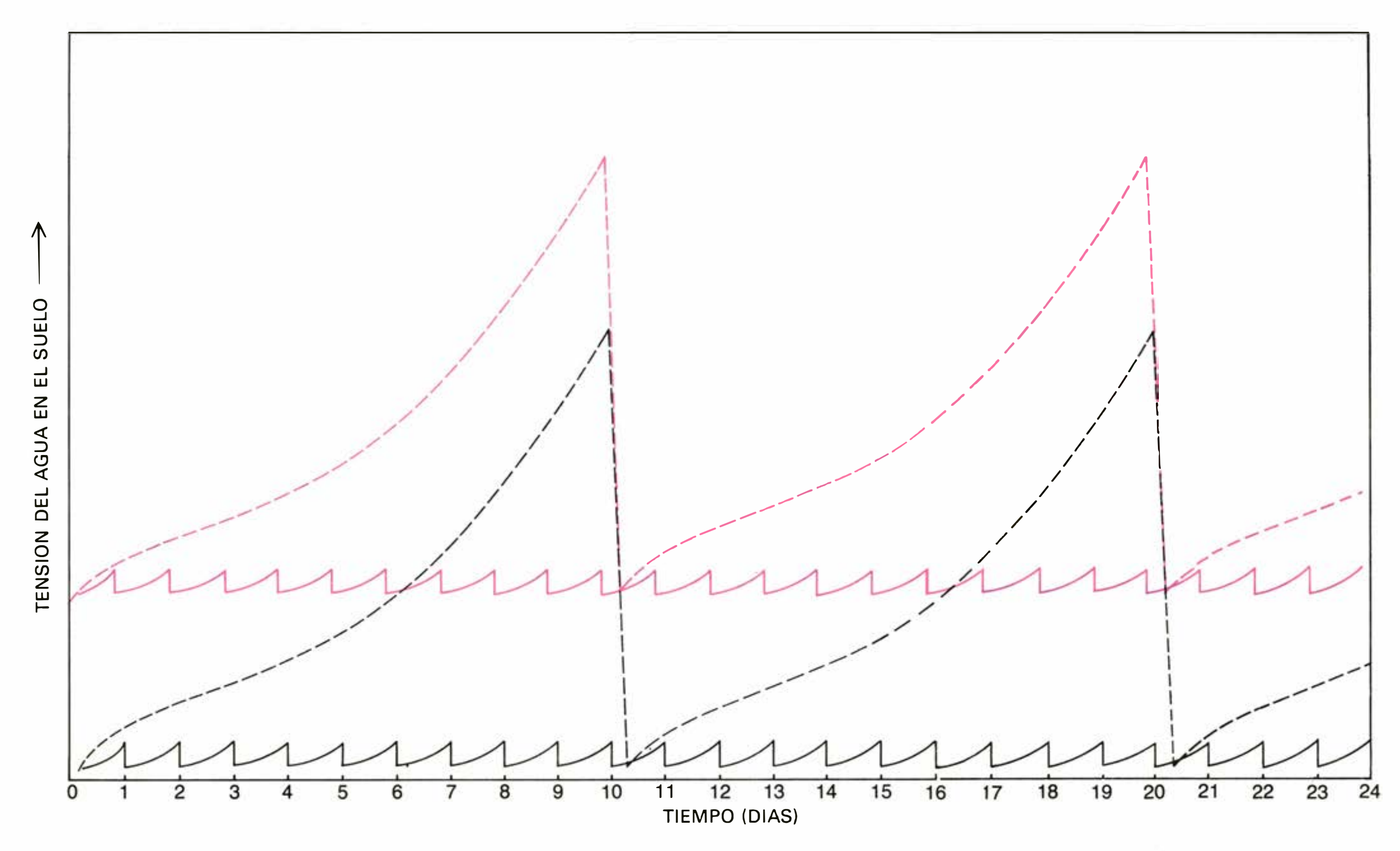
El sistema requiere también un puesto de control al pie del campo. Típicamente, este puesto contiene dispositivos para filtrar el agua y eliminar las partículas que puedan obturar las líneas laterales, un inyector que permita incorporar abono o productos químicos al agua, reguladores de presión, contador de agua y, por supuesto, las válvulas y bombas necesarias para regular el flujo del agua. Suele tratarse de un sistema diseñado de manera que pueda accionarse manual o automáticamente (mediante relojes y válvulas). También puede diseñarse para regar un determinado número de campos en el orden deseado.

Aunque el número de hectáreas con el riego gota a gota, en Estados Unidos, ha aumentado desde 40 en 1960 a más de 54.600 en la actualidad, el total resulta todavía pequeño si lo comparamos con la extensión regada mediante otros métodos. La tendencia, sin embargo, es de fuerte expansión. Según un estudio sobre el riego gota a gota, dirigido en 1975 por C. Don Gustafson, de la University of California Agricultural Extension, California tenía la mayor extensión de suelo (24.290 hectáreas) regadas por los métodos de gota a gota, seguido a distancia (aunque en proporciones significativas) por Texas, Hawaii, Florida, Arizona y Michigan. Los sistemas gota a gota se han instalado también ampliamente en Australia, Israel, Africa del Sur y México, y, en menor cantidad, en Canadá, Chipre, Francia, Irán, Nueva Zelanda, Gran Bretaña y ciertas zonas de Africa occidental. Se calcula que, en todo el mundo, habrá unas 350.000 hectáreas que utilizarán el riego gota a gota en 1980.

Lo que hace económicamente atractivo al riego gota a gota en muchas situaciones agrícolas es que proporciona más beneficios que otros métodos. Probablemente, la principal ventaja que no se obtiene en otras técnicas sea que el riego gota a gota abastece a las plantas de la cantidad de agua precisa. La cantidad justa necesaria llega a la zona radicular de la planta para restituir la cantidad consumida en la evapotranspiración (el agua evaporada del suelo y la transpirada por la planta), así como una cantidad adicional para lavar las sales de la zona del suelo inmediata a las raíces. Este procedimiento es mucho más eficaz que inundar todo el campo o suministrar agua por surcos.

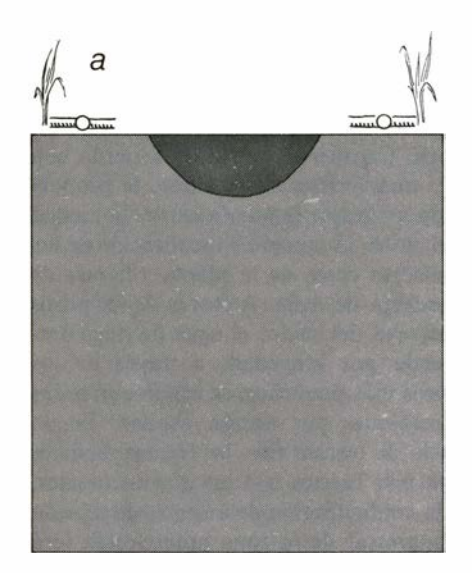
En el pasado se trabajó con ahínco para hallar las ecuaciones que representasen la cantidad de agua necesitada por las distintas especies de plantas. Los factores que integran la ecuación abarcan: la naturaleza de la superficie de evaporación y los efectos del viento y la temperatura sobre la presión de vapor, la calidad del agua y la cantidad de energía disponible. Una vez determinadas las exigencias de agua para un cultivo se programa, mediante la instalación del riego gota a gota, la de cada planta. A medida que el agua fluye desde los emisores se introduce bajo tierra, creando una estructura semejante a la de un "recipiente" de humedad alrededor de las raíces de la planta.

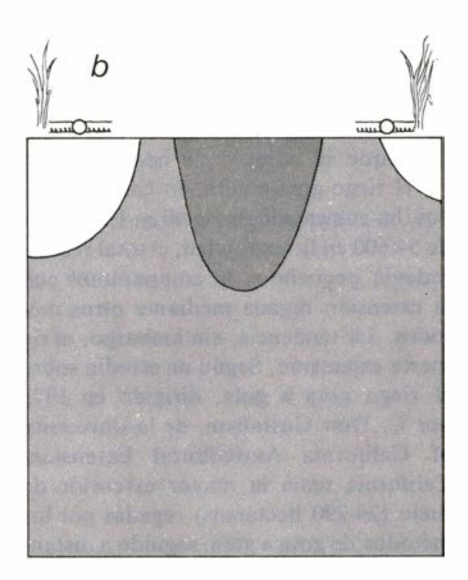
El tamaño y la forma de la zona humedecida, como podría observarse en un corte transversal, varía de acuerdo con las características del suelo, la proporción en la que la planta extrae humedad del suelo, el número y localización de los emisores cerca de la planta y la tasa de descarga de agua. A través de los poros mayores del suelo, el agua de riego desciende por gravedad; a través de los poros más pequeños, se esparce en todas direcciones por acción capilar. En un suelo de textura fina, las fuerzas capilares son más fuertes que las gravitacionales, y la configuración de un corte de sección transversal de la zona humedecida será más o menos circular. En un suelo de textura arenosa, que no es adecuado para la buena retención del agua, la sección transversal adquirirá una configuración más elíptica.

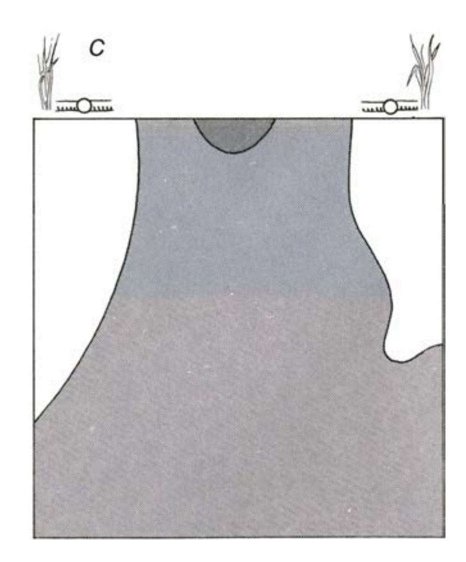


TENSION DE HUMEDAD DEL SUELO, que mide la fuerza con que el agua queda retenida en el mismo y, por tanto, el esfuerzo realizado por una planta para extraer el agua, comparada en el riego gota a gota *(curvas continuas)*, con el riego por aspersión y en surcos *(curvas discontinuas)*. En el riego gota a gota, el agua se aplica diariamente. La tensión de humedad del suelo sube ligeramente a lo largo del período de 12 a 18 horas, entre aplicaciones. Con el riego por aspersión o en surcos, que suele darse a intervalos de tiempo mayores, la tensión de humedad sube notablemente entre aplicaciones. El conjunto superior de las curvas representan el riego con agua salina; el inferior, el riego con agua dulce.

Debido a que el riego gota a gota suministra agua lenta y frecuentemente a una tasa prestablecida, el contenido de humedad en la zona mojada permanece aproximadamente constante. Así pues, la planta crece sin tensión en un medio cuya humedad es la idónea. El riego gota a gota elimina de forma eficaz las amplias fluctuaciones en el contenido de humedad del suelo que resultan con frecuencia cuando el agua se aplica periódicamente por otros métodos de riego. Por ejemplo, entre aplicaciones de agua por aspersión o mediante surcos, parte de ella se pierde por evaporación del suelo y por transpiración de la planta, y la que queda se le hace a la raíz cada vez más difícil de extraer. La tensión de humedad del suelo (la fuerza con que el agua está retenida a la matriz del suelo) aumenta; si se prolonga







DESARROLLO DE CONCENTRACION DE SALES en la zona radicular de plantas sometidas al riego gota a gota; en los tres casos, las plantas están separadas un metro, y cada ejemplar tiene cerca de su base un emisor. Se representan cuatro grados de salinidad, que varían desde insignificantes a graves, mediante un espectro que oscila desde blanco, pasando por tonos grises, hasta negro. Si la cantidad de agua de riego es menor que la cantidad perdida por evaporación y transpiración *(a)*, se desarrolla una concentración elevada de sales entre las plantas, y toda la zona radicular contiene una alta concentración de sales. En los otros dos casos, el riego compensa *(b)* o supera *(c)* a la evapotranspiración.

indebidamente a nivel alto, las plantas sufren tensiones y pueden comenzar a marchitarse.

Con el riego gota a gota, sin embargo, un cultivador puede programar las aplicaciones de agua para mantener un margen estrecho y predeterminado de los niveles de tensión de humedad del suelo. El nivel óptimo se denomina capacidad de campo; se trata del nivel en que el agua excedente se drena del suelo por gravedad y el agua que permanece queda retenida en el suelo por acción capilar. En la capacidad de campo, cuando la tensión de humedad del suelo va de 0 a 0,3 atmósferas, las plantas pueden absorber agua del suelo con un esfuerzo mínimo. Puesto que las exigencias de agua de un cultivo en un campo responden fundamentalmente a un cálculo aproximativo, no obstante las ecuaciones obtenidas para la evapotranspiración, se disponen a menudo instrumentos de medición en el campo para controlar el contenido de humedad en la zona radicular, con el fin de asegurar que éste se acerque al máximo a la capacidad del campo.

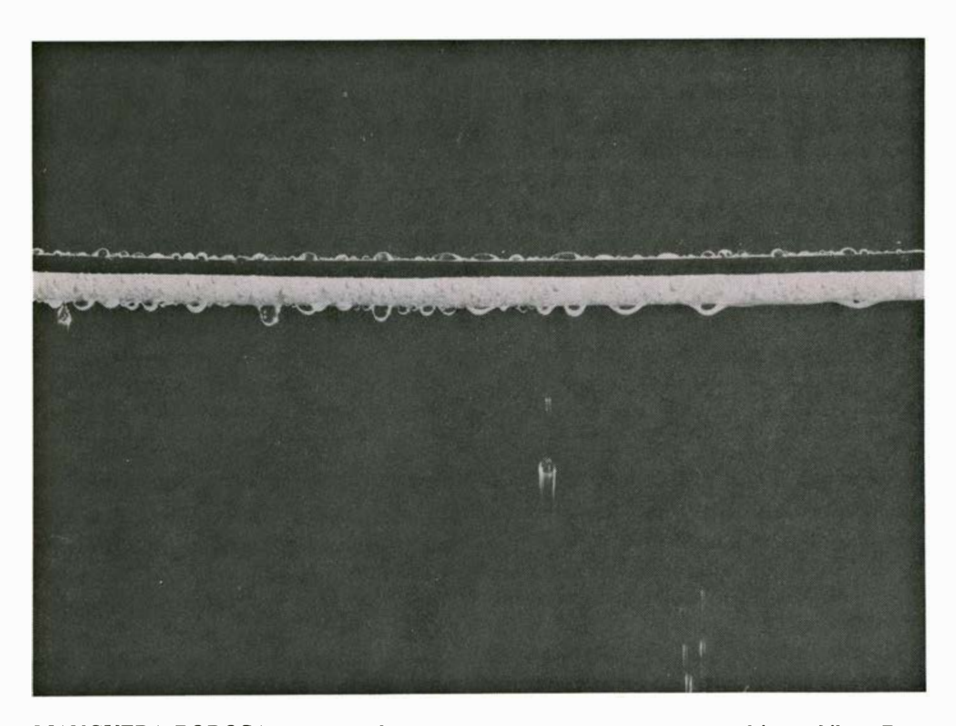
La idea de mantener la tensión de humedad del suelo a nivel óptimo para el crecimiento de un cultivo no es nueva. Sin embargo, antes del desarrollo del riego gota a gota resultaba económicamente impracticable intentar mantener tal condición con los métodos de riego tradicionales, habida cuenta de las desmesuradas cantidades de agua y el enorme trabajo necesarios. Una ventaja que resulta de la eliminación de la tensión en la planta por el riego gota a gota es que las plantas suelen alcanzar la madurez antes que otras plantas regadas por métodos distintos. Una cosecha temprana puede venderse a menudo a precios más elevados que una cosecha puesta en el mercado en plena temporada.

Otra ventaja del riego gota a gota es la facilidad para sacar el máximo rendimiento del agua disponible. Esta ventaja ha despertado el interés de las zonas de todo el mundo que padecen escasez de agua.

Con el riego gota a gota, la porción de terreno que media entre dos filas de plantas permanecen secas, de suerte que se pierde poca agua por evaporación, escorrentía o percolación. (A veces, en el riego por aspersión, aparece una forma adicional de pérdida al dispersarse el agua por el viento antes de llegar a las plantas.) Las plantas gastan tanta agua en el riego gota a gota como con otros métodos, pero con la ventaja de que allí

PAIS	HECTAREAS EN 1975	CULTIVOS
ARGENTINA	150	VIÑEDOS
AUSTRALIA	17.136	FRUTALES DE HOJA CADUCA, HORTALIZAS, AGUACATES, BANANOS, AGRIOS, PLANTAS DE VIVERO, NOGALES
BRASIL	1364	FRUTALES DE HOJA CADUCA, HORTALIZAS, PLANTAS DE VIVERO, NOGALES
CANADA	400	FRUTALES DE HOJA CADUCA, HORTALIZAS, PLANTAS DE VIVERO, ORNAMENTALES
CHIPRE	400	HORTALIZAS, VIÑAS, BANANOS, AGRIOS
COSTA RICA	10	PALMAS DE GUINEA (EXPERIMENTAL)
FRANCIA	1220	FRUTALES DE HOJA CADUCA, FLORES, PLANTAS DE INVERNADERO
HONDURAS	2	BANANOS, PALMERAS
INDIA	20	FRUTALES DE HOJA CADUCA, HORTALIZAS
IRAN	800	AGRIOS, HORTALIZAS
ISRAEL	10.000	FRUTALES DE HOJA CADUCA, HORTALIZAS, AGUACATES, BANANOS, AGRIOS, PLANTAS DE INVERNADERO, ORNAMENTALES
JAPON	4	AGRIOS, PARRAS, HORTALIZAS
MEXICO	5530	FRUTALES DE HOJA CADUCA, AGRIOS, CULTIVOS EN HILERA, OLIVOS
MARTINICA	400	BANANOS, AGUACATES, CAÑA DE AZUCAR
NUEVA ZELANDA	1073	FRUTALES DE HOJA CADUCA, HORTALIZAS, AGRIOS, AGUACATES, PLANTAS DE INVERNADERO
PANAMA	0,8	FRUTALES DE HOJA CADUCA, HORTALIZAS
PUERTO RICO	72	MANGOS, CEREZOS, PLATANEROS
SUDAFRICA	7200	FRUTALES DE HOJA CADUCA, HORTALIZAS, FLORES DE INVERNADERO
GRAN BRETAÑA	1600	FRUTALES DE HOJA CADUCA, HORTALIZAS, FLORES DE INVERNADERO
ESTADOS UNIDOS	53.487	FRUTALES DE HOJA CADUCA, HORTALIZAS, AGRIOS, CAÑA DE AZUCAR, NOGALES, PLANTAS DE INVERNADERO, PLANTAS DE VIVERO, ORNAMENTALES, CULTIVOS EN HILERA, AGUACATES
AFRICA OCCID. (SENEGAL)	400	HORTALIZAS

EXTENSION DEL RIEGO GOTA A GOTA, según datos de 1975. Los datos están basados en un estudio realizado por C. Don Gustafson, de la University of California Agricultural Extension. En Estados Unidos, la mayor extensión de campos (24.000 hectáreas) que empleaban el riego gota a gota correspondía al estado de California. Otros siete estados regaban más de 400 hectáreas con ese mismo método inovador. C. Don Gustafson encontró ese sistema de riego en 35 estados.



MANGUERA POROSA, que gotea lentamente para regar gota a gota un cultivo en hilera. Para obtener esta fotografía, se ha sacado la manguera de la tierra con el fin de resaltar claramente la emisión del agua. La franja negra, a lo largo de la parte superior del tubo, es una lacra de resina.

las pérdidas se reducen al mínimo. La eficacia del uso del agua varía con la buena administración del sistema gota a gota. Los cosechadores de caña de azúcar de Hawaii han informado de incrementos en la eficacia del 40 al 60 por ciento mediante un empleo adecuado del riego gota a gota.

En Hawaii, que suministra cerca del 10 por ciento del azúcar consumido en Estados Unidos, se riegan unas 48.600 hectáreas de las 89.700 totales destinadas a plantaciones de caña de azúcar. Cada hectárea regada requiere 46.700 metros cúbicos de agua a lo largo de los dos años del período de desarrollo de la caña. Ya que la disponibilidad de agua dulce en Hawaii es muy limitada y el valor de la misma normalmente alto, el uso eficaz del agua de riego resulta esencial. La eficacia en el riego se mide por la relación entre la cantidad de agua disponible para la planta y la cantidad aplicada. Los sistemas tradicionales de riego por surcos en Hawaii son muy ineficaces, al consumirse sólo la mitad del agua aplicada, aproximadamente. Los sistemas por aspersión se muestran más eficaces, con una relación de eficacia del 70 al 80 por ciento, pero el coste de la operación y los bienes de equipo son elevados. Con el riego gota a gota se obtienen rendimientos entre el 80 y 95 por ciento.

El tercer beneficio importante del riego gota a gota es que funciona bien aun cuando el agua tenga una concentración salina alta. La mayoría de los cultivos de regadío podrán tolerar agua con un contenido total de sales disueltas de hasta 600 miligramos por litro. Si el drenaje y lavado son los adecuados, el riego cumple su misión con agua que contenga desde 500 a 1500 miligramos de sales disuelta por cada litro de agua. Si la salinidad está entre 1000 y 2000 miligramos por litro, el riego debe ser frecuente para promover la lixiviación. El agua cuya salinidad oscile entre 3000 y 5000 miligramos por litro, sólo proporcionará una cosecha productiva cuando las plantas que se cultiven sean muy tolerantes a las sales.

Al regarse con agua salina, aumenta la concentración de sales en el suelo a medida que éste se va secando entre las aplicaciones de agua. En ese período de tiempo, la tensión de humedad del suelo aumenta, lo cual dificulta que las plantas extraigan del suelo el agua de riego remanente. Las sales se van acumulando gradualmente, y disminuye el crecimiento de las plantas y el rendimiento de la cosecha. Entre agricultores es costumbre confiar en la lluvia o en las fuertes aplicaciones adicionales de agua de riego para lavar de sales el suelo. Sin embargo, sólo con una regulación adecuada de las relaciones entre la planta, el suelo y el agua puede mantenerse una buena productividad, en particular en las regiones áridas. A veces, ni siquiera con una adecuada regulación pueden obtenerse buenas producciones cuando se riega con agua salina con los métodos convencionales.

Por otra parte, con el riego gota a gota, se controla el incremento de las sales, al someter el suelo a una lixiviación eficaz y continua. Las sales son empujadas hacia la periferia del perfil de la raíz por un frente de agua que avanza desde los orificios del tubo. Las raíces pueden tomar agua libremente en medio de la zona húmeda, en donde la tensión de humedad del suelo es baja y el nivel de las sales permanece aproximadamente como en el agua de riego.

El riego gota a gota suele comportar un rendimiento mayor y un crecimiento más uniforme de la cosecha. El creci-

AGUJERO DE EMISION
ENTRADA DEL AGUA
ENTRADA A LA TRAYECTORIA ESPIRAL

EMISOR TIPO BOQUILLA, diseñado para menguar la presión del agua, al obligarle a pasar por una trayectoria espiral alargada. A presión baja, el agujero del emisor puede ser mayor, y así se frena su tendencia a obturarse. En la fábrica se fuerza a un emisor de este tipo a emitir 3,8, 5,7 o 7,6 litros de agua por hora. El gancho de la derecha fija el emisor a la tubería de riego gota a gota en un ángulo recto; el agua va penetrando, desde el centro de la tubería, con inclinación cónica.

miento uniforme, que resulta de la aplicación de la misma cantidad de agua, aproximadamente, a cada planta del campo, facilita en gran manera la recolección mecánica. El riego gota a gota, como ya he mencionado, no exige ninguna preparación especial del suelo. Además, el sistema se hace adaptable a una amplia variedad de tipos de suelos y campos. Puesto que la humedad puede aplicarse de forma continua, el que ciertos suelos (los arenosos, por ejemplo) no retuvieran agua entre ciclos de riego carece de importancia. Al eliminar el problema de la escorrentía, el riego gota a gota puede instalarse en zonas montañosas. Con el riego gota a gota, la caña de azúcar se planta en Hawaii en pendientes del 20 al 30 por ciento, y existen campos de aguacates en San Diego en pendientes del 50 al 60 por ciento, sin problemas de erosión.

El riego gota a gota saca también el máximo provecho de los fertilizantes. Mediante la aplicación directa de abonos solubles a través de la instalación del riego gota a gota, no sólo se obtiene una liberación de los nutrientes más perfecta y constante, sino que se reducen asimismo los costes al eliminar las pérdidas de fertilizante por percolación y entre surcos. El sistema reduce los problemas ambientales que van asociados con la contaminación de aguas subterráneas por productos químicos agrícolas. Además, las plantas no experimentan nunca quemaduras por abono en un sistema gota a gota, puesto que los productos químicos del abono se diluyen extensamente en el agua de riego antes de alcanzar la planta.

Al permanecer seca con el riego gota a gota la mayoría de la superficie del suelo, se inhibe el crecimiento de malas hierbas y de hongos. La maleza puede convertirse en un problema en las zonas humedecidas cercanas a los emisores. Ciertos experimentos llevados a cabo en California han señalado que esas malas hierbas pueden controlarse satisfactoriamente con herbicidas.

Una última ventaja del sistema es que puede reducir significativamente la cantidad de dinero invertido en energía y equipamiento. Al pie del campo, la presión de un sistema de riego gota a gota típico es de 0,4 kilogramos a un kilogramo por centímetro cuadrado, mientras que con el riego por aspersión varía entre tres y ocho kilogramos por centímetro cuadrado. Se necesita menos potencia para el bombeo en el riego gota a gota, y la baja presión del sistema recorta también el gasto de material, porque las cañerías principales y las secundarias pueden ser de una sección menor que las usadas en el riego por aspersión.

El riego gota a gota no está exento de problemas. Quizás el más grave sea la obturación de agujeros y emisores, que pueden dañar la eficacia de la distribución a través del sistema. Las bajas presiones, los orificios pequeños y la velocidad baja de flujo aumentan desmesuradamente la obturación que puede producirse por partículas transportadas por el agua o por el lodo que se concentra en torno los orificios o revisten el interior de los tubos.

La investigación demuestra que las partículas pueden tratarse con filtros. Las cubetas de sedimentación pueden extraer partículas cuyo diámetro supere las 75 micras. Las partículas menores pueden extraerse con filtros de criba o, en casos graves, con filtros de arena. La obturación puede reducirse con ciertas prácticas de campo, tales como la orientación de los orificios hacia arriba y mediante el lavaje periódico de los tubos. El problema del lodo puede eliminarse recurriendo a la cloración del agua durante 20 minutos por día, a razón de 10 partes por millón. El cloro sirve, simultáneamente, para controlar las bacterias que prosperan en el hierro y azufre de aguas de ciertas zonas. Otro problema que a veces va asociado al riego gota a gota consiste en el ataque que sufren los tubos por parte de hormigas en busca de agua y roedores que afilan sus dientes. El método que se sigue hoy para hacer frente a tales ataques estriba en aplicar productos químicos. La investigación está centrada en la incorporación de repelentes en los materiales de que están construidos los tubos.

En la Conferencia Mundial del Agua (convocada por las Naciones Unidas, en marzo del año pasado, en Argentina) se señaló que aunque el 70 por ciento de la superficie de la tierra está cubierta por agua, sólo el 1 por ciento del agua es dulce y que el 99 por ciento del agua dulce es subterránea. En zonas que sufren escasez de agua, a medida que se encarecen los costes de la construcción de pozos excavados, los desvíos de ríos, la desalinización del agua de mar y la formación de nubes, la conservación y utilización eficaz de las fuentes disponibles se hacen cada vez más importantes. El riego gota a gota es una técnica prometedora para ayudar a solucionar la escasez de agua y alimentos, al mejorar la eficacia del riego e incrementar la productividad de la tierra.

El agrupamiento de las galaxias

Las galaxias tienden a formar grupos pequeños, que, a su vez, constituyen cúmulos mayores, y así sucesivamente. Tal organización jerárquica, por largo tiempo sospechada, no pudo verse con claridad hasta hace poco

Edward J. Groth, P. James E. Peebles, Michael Seldner y Raymond M. Soneira

La cosmología moderna se basa en la noción de que la materia está uniformemente distribuida a lo largo de todo el universo. Este principio está apoyado por muchas pruebas observacionales; pero, por otra parte, esto no puede ser la verdad absoluta. En nuestro propio entorno, la materia se distribuye de una forma totalmente irregular: se presenta muy densa en el interior de las estrellas y muy rarificada en el espacio interestelar. Las estrellas se concentran a su vez en galaxias y están prácticamente ausentes en los enormes volúmenes intergalácticos. Más aún, ni siquiera la distribución de las galaxias es uniforme. Las galaxias se organizan en cúmulos y éstos tienden a formar supercúmulos. En consecuencia, las estructuras astronómicas guardan una jerarquía, y sólo un cálculo promediado sobre varios miles de galaxias da lugar a una distribución de la materia que se aproxime a la uniformidad.

Recientemente, hemos podido demostrar que la agrupación de galaxias presenta una pauta fundamental de notable sencillez y regularidad. El descubrimiento de tales pautas de comportamiento en la naturaleza es siempre excitante, ya que suele comportar un aumento de nuestra comprensión sobre la misma. En este caso, la pauta no sólo describe la actual estructura a gran escala del universo sino que impone también limitaciones a los modelos que intentan describir cómo pudo haber comenzado el universo. La jerarquía de los cúmulos de galaxias que observamos hoy es la jerarquía que debiera esperarse que se diera en un universo que haya evolucionado desde un estado primitivo de alta densidad, en otras palabras, desde la "gran explosión".

Nuestra propia galaxia tiene una notable compañera próxima, la Gran Nebulosa de Andrómeda. Dentro de un radio de alrededor de 2 millones de años-luz hay varias docenas de pequeñas galaxias, todas las cuales están probablemente ligadas gravitacionalmente unas a otras. Estas galaxias, junto con la nuestra, reciben el nombre global de el Grupo Local. Un cúmulo similar, llamado el grupo M81 en razón de su miembro más caracterizado, aparece relativamente próximo. El grupo M81 proporciona una excelente ilustración de una regla empírica e informal: el mejor lugar para localizar una galaxia es justo al lado de otra galaxia. Pocas galaxias se hallan tan cercanas a nosotros como lo está M81; y donde encontramos esa galaxia, localizamos, como mínimo, otras dos más en el mismo entorno.

El Supercúmulo Local

Un conjunto mucho mayor de galaxias, el cúmulo de Virgen, se halla centrado en torno a una posición que se encuentra a unos 50 millones de años-luz de nosotros. Una nube de galaxias, bastante densa se extiende desde el cúmulo de Virgen hasta aproximadamente nuestra propia posición. Por tanto, un observador situado en un punto alejado del universo vería inmediatamente que nuestra galaxia es como la frontera de una concentración de varios miles de galaxias centrados en el cúmulo de Virgen. El sistema en conjunto recibe el nombre de Supercúmulo Local.

Los estudios sobre agregación de galaxias comienzan generalmente con la preparación de mapas de todas las galaxias que en una región del cielo exceden cierto nivel convencional de brillo aparente. Este enfoque del problema presenta varias limitaciones serias: las galaxias están distribuidas en un espacio tridimensional, pero en un mapa su ordenación se reduce a un plano bidimensional. Dos galaxias que sean adyacentes en tal mapa pueden hallarse, en realidad, a distancias notablemente diferentes de cuanto las contemplamos a lo largo de la línea de visión, y, por consiguiente, no estarían próximas en absoluto. Se pueden estimar las distancias a que se hallan las galaxias; pero cuando deben calcularse las posiciones de miles de galaxias el tiempo necesario para realizar las mediciones se prolonga enormemente y el método se hace impracticable.

Podemos, sin embargo, eludir el problema del desconocimiento de las distancias cuando se ha realizado un mapa con gran número de galaxias; en realidad, las distancias pueden ignorarse a condición de que la muestra de galaxias sea grande. La separación angular en un mapa del cielo no es una indicación fidedigna de la distancia entre un determinado par de galaxias, pero sí es significativa cuando se promedia sobre miles de parejas.

Los mapas de las distribuciones de galaxias se basan en el brillo aparente, que se expresa usualmente en términos de magnitud. Cuanto mayor es la magnitud de un objeto astronómico más débilmente se manifiesta. La escala de magnitudes es una escala logarítmica, calibrada de forma que un incremento de cinco magnitudes corresponda a una disminución en el brillo de cien veces. Las estrellas más brillantes son de primera magnitud; los objetos más débiles que podemos ver con los ojos, sin auxilio alguno, son aproximadamente de sexta magnitud. Debido a que las galaxias son objetos lejanos, la mayoría de ellas aparecen con extremada debilidad y, por tanto, presentan altas magnitudes.

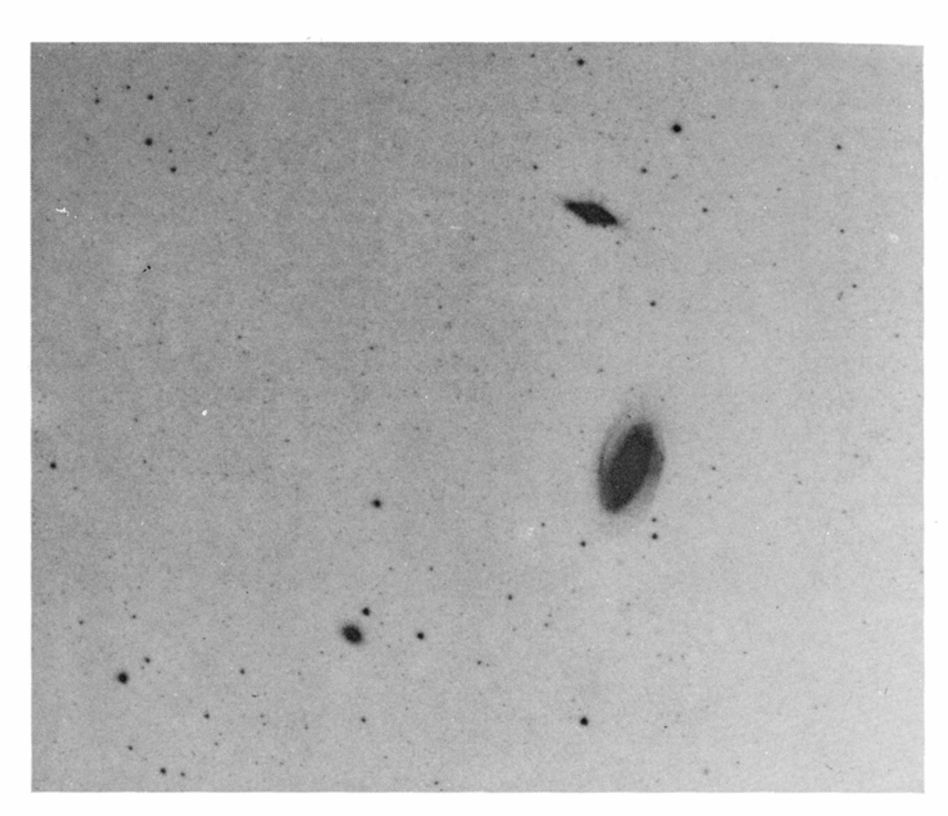
Los datos coleccionados por el desaparecido Fritz Zwicky y sus colegas, del Instituto de Tecnología de California, pueden presentarse en forma de mapas de galaxias que cubren la mayor parte

del hemisferio norte celeste. Uno de tales mapas comprende galaxias más brillantes que las de magnitud 13, lo cual representa un brillo 600 veces más débil que el de los objetos menos brillantes que puede ver el ojo sin ayuda. El mapa recoge unas 500 galaxias con una distancia promedio de unos 140 millones de años-luz. El cúmulo de Virgen aparece en este mapa como un grupo denso; además, puede percibirse la banda de galaxias que se extiende desde el cúmulo de Virgen hasta nuestra galaxia.

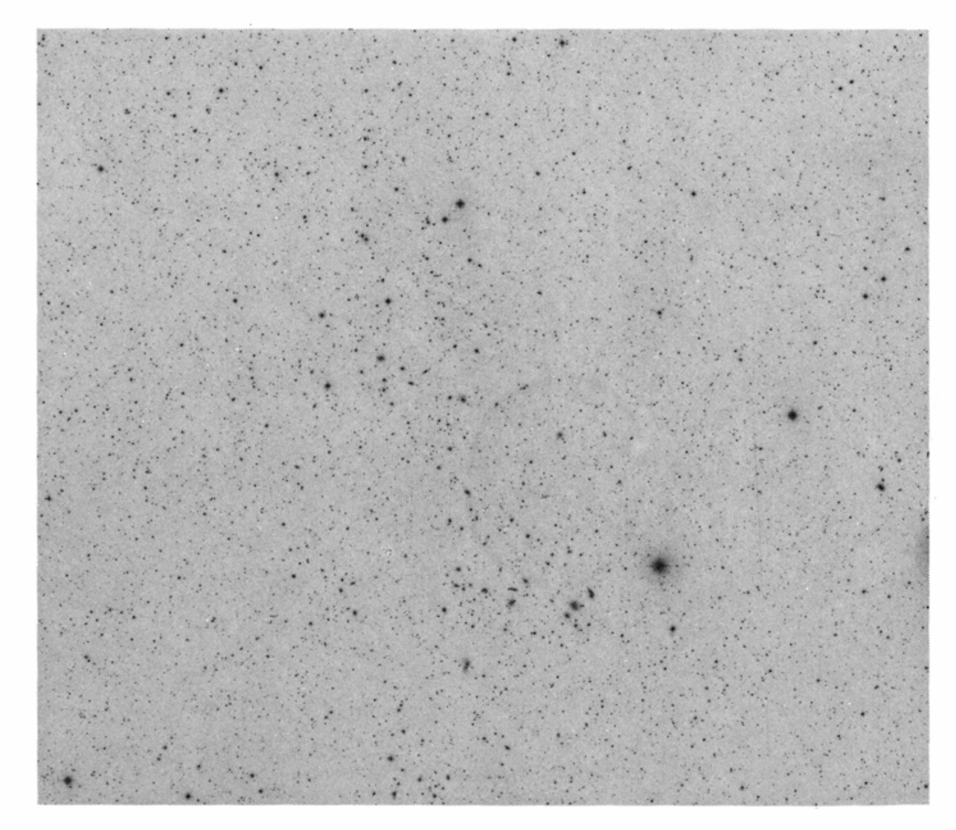
Otro mapa basado en los datos de Zwicky abarca galaxias más brillantes que las de magnitud 15, unas 4000 veces más débiles que el límite de percepción del ojo. Para tal brillo aparente se pueden distinguir más de 5000 galaxias, con unas distancias típicas de 340 millones de años-luz. El Supercúmulo Local comprende ahora muchas más galaxias, debido simplemente a que se tienen en cuenta miembros del supercúmulo más débiles. Lo que es más, la distribución de galaxias en cualquier parte del cielo resulta más uniforme para los mapas que tienen en cuenta un número mayor de galaxias. El agrupamiento no es menos pronunciado para grandes distancias; la distribución parece más regular, por la sencilla razón de que muchos cúmulos se solapan a lo largo de la línea de visión.

Una colección más extensa de galaxias ha sido compilada por C. Donald Shane y Carl A. Wirtanen, del Lick Observatory. Su mapa cubre esencialmente la misma zona del cielo que la muestra de Zwicky, pero abarca galaxias más brillantes que las de decimonovena magnitud; un objeto de decimonovena magnitud es del orden de 160.000 veces más débil que otro escasamente visible a simple vista. Las galaxias se hallan a una distancia promedio de 1400 millones de años-luz, lo que representa una lejanía 4,4 veces superior a la de los objetos de decimoquinta magnitud investigados por Zwicky. El volumen del espacio estudiado es $(4,4)^3$, aproximadamente 85, veces mayor. Como era de esperar, si la densidad de la materia es la misma en todo el universo, el mapa de Shane-Wirtanen contiene 85 veces más galaxias que el mapa de Zwicky, esto es, aproximadamente un millón en total. La recopilación de todas las galaxias necesitó 12 años y se realizó con un cuidado tan meticuloso que el mapa Shane-Wirtanen se ha convertido en una fuente indispensable de información sobre la agrupación de galaxias.

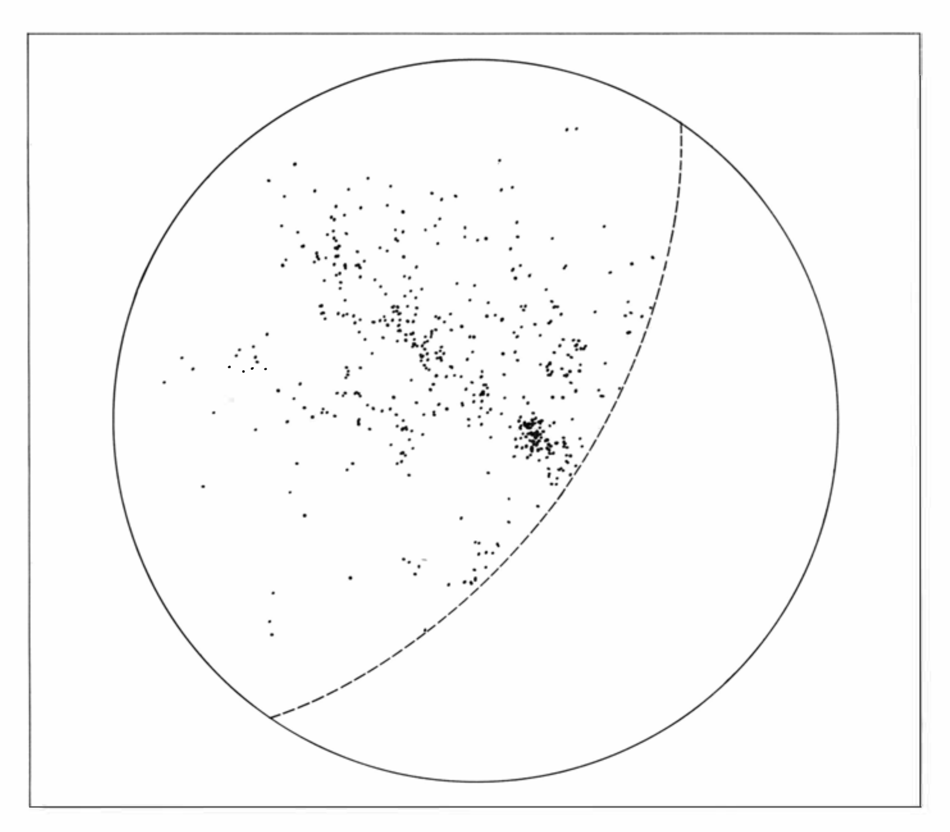
Contar un millón de objetos, aunque se empeñen 12 años, no permite detenerse mucho tiempo en cada uno. Shane y Wirtanen dividieron el cielo en celdas



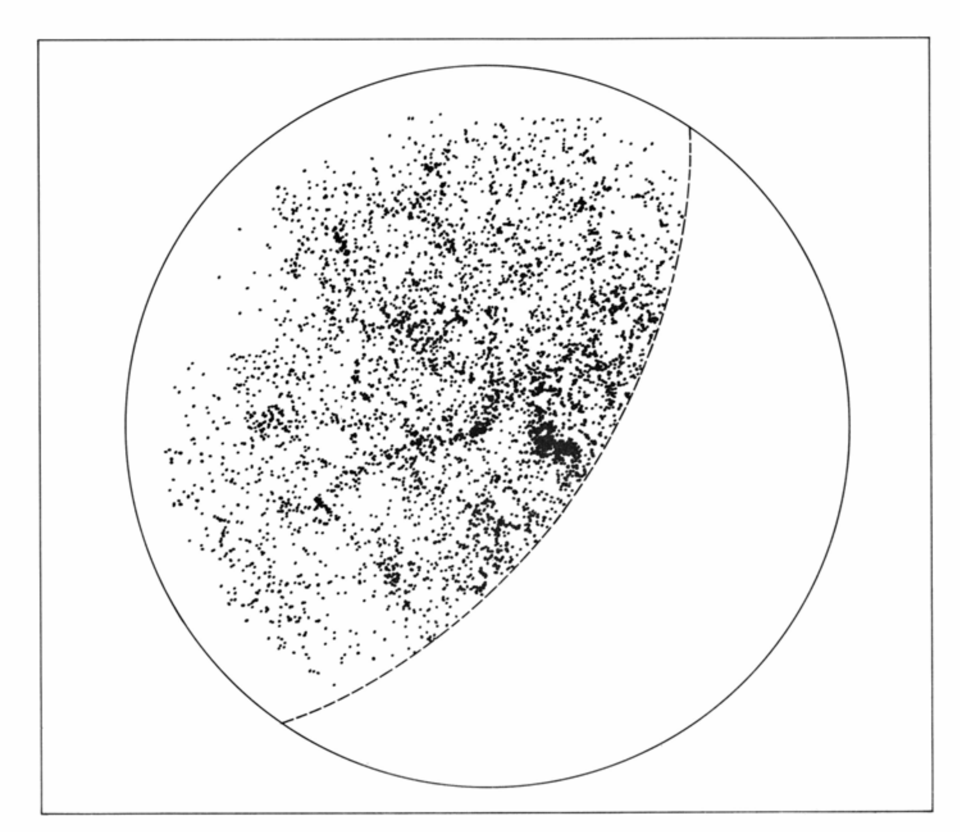
CUMULO DE GALAXIAS próximo conocido por grupo M81 en honor de su miembro más destacado, que figura con el número 81 en el catálogo de galaxias compilado por Charles-Joseph Messier. M81 es la galaxia mayor, a la derecha del centro de la figura. Es muy semejante a nuestra propia galaxia y el grupo M81 es similar al Grupo Local, el cúmulo que incluye a nuestra galaxia. M81 se encuentra a una distancia de nosotros de unos 10 millones de años-luz. Esta fotografía y la inferior se han tomado con el telescopio Schmidt de 48 pulgadas de Monte Palomar (Estados Unidos).



GRAN CUMULO de galaxias en la constelación de Hércules, a una distancia de unos 700 millones de años-luz. En la fotografía pueden identificarse unas 20 o 30 galaxias, pero el cúmulo comprende también cientos de galaxias más débiles que sólo pueden observarse con largas exposiciones. Las imágenes de las galaxias pueden distinguirse de las que corresponden a estrellas por la ligera nebulosidad que rodea a las primeras. En la jerarquía de las estructuras astronómicas, el cúmulo de Hércules es una entidad mucho mayor que el grupo M81; para una distancia de unos 700 millones de años-luz, la sección del cielo que se muestra en esta fotografía mide en torno a los 20 millones de años-luz.



LA DISTRIBUCION DE GALAXIAS en el hemisferio norte celeste se cartografió a partir de datos compilados por Fritz Zwicky, del Instituto de Tecnología de California. Cada punto representa una galaxia; en total hay casi 500. El mapa comprende galaxias que exceden la magnitud 13 en brillo aparente. Las galaxias tienen una distancia promedio de 140 millones de años-luz. La agrupación de galaxias próxima al centro del mapa es el cúmulo de Virgen; la banda que se extiende hacia la posición que ocuparía las 10 horas si el círculo representase una esfera de reloj, conecta el cúmulo de Virgen con el Grupo Local, y así constituyen el Supercúmulo Local.



LA MAYOR MUESTRA de galaxias elaborada por Zwicky abarcó las que superan la magnitud 15. En el mapa hay más de 5000 galaxias, con una distancia típica de 340 millones de años-luz. El cúmulo de Virgen aparece mucho más denso, pero las características del Supercúmulo Local quedan difuminadas debido a la superposición de cúmulos que presentan distancias radiales distintas.

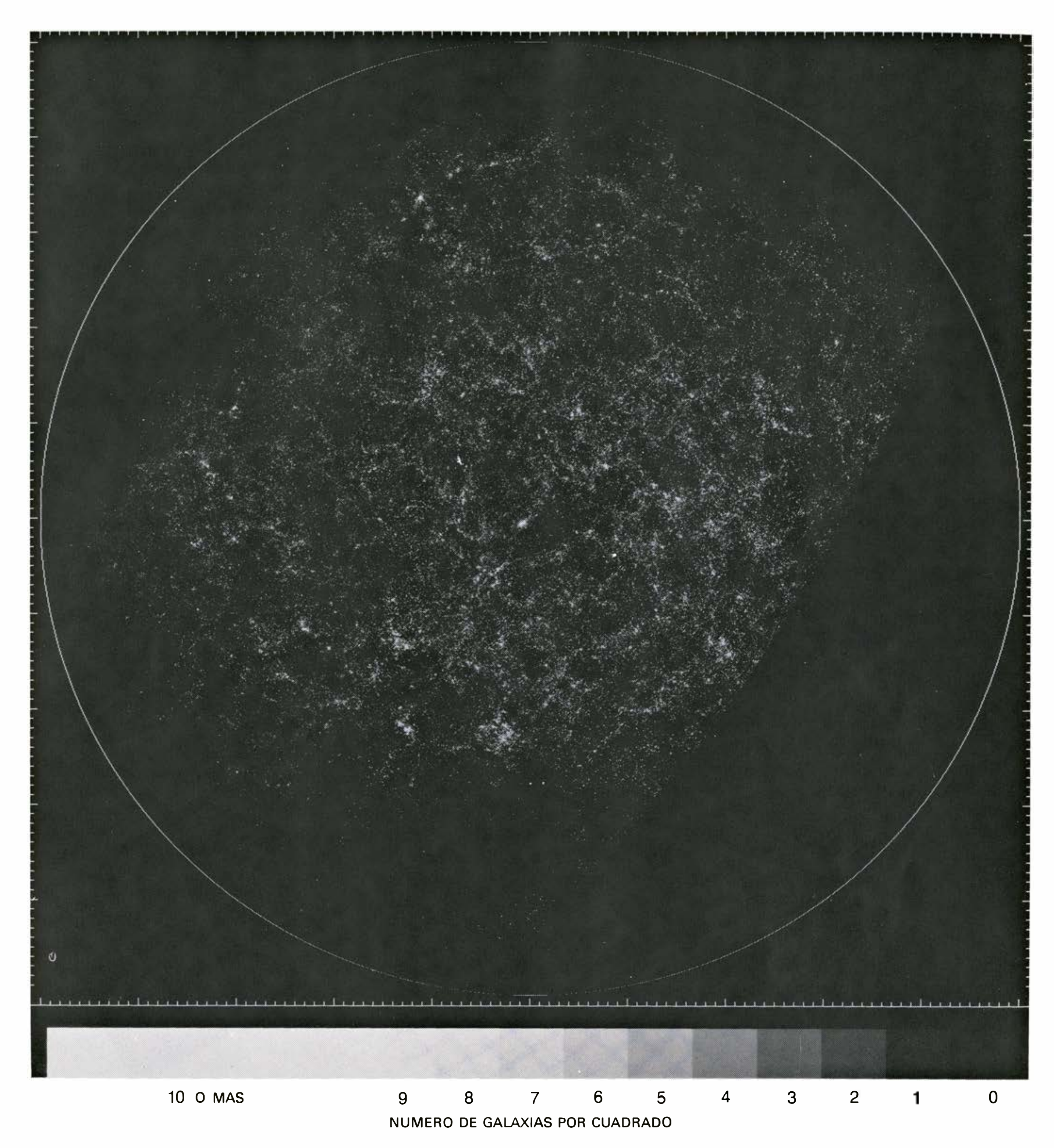
de un sexto de grado por lado y contaron, en cada celda, el número de galaxias que brillaban más que las de decimonovena magnitud. Examinaron un millón de dichas celdas en placas fotográficas. En promedio, se presentaba una galaxia por celda, pero se registraron variaciones significativas; por ejemplo, se encontraron 1600 celdas que contenían 10 o más galaxias.

Medidas estadísticas

Las galaxias del Supercúmulo Local están incluidas en el mapa de Shane-Wirtanen, pero hay tal cantidad de otras galaxias a mayores distancias que ese rasgo de vecindad se pierde entre la multitud. Otros cúmulos, sin embargo, son distinguibles. En la constelación Coma Berenices aparece una colección densa y elíptica de galaxias que recibe el nombre de cúmulo de Coma. Es unas 6 veces más distante que el cúmulo de Virgen, y contiene 10 veces más galaxias. Un rasgo más general del mapa de Shane-Wirtanen es la curiosa estructura filamentosa que parece cubrir todo el cielo. La interpretación de los filamentos es un tanto incierta.

Ultimamente ha podido confeccionarse un mapa de galaxias, por lo menos, que se extiende a mayor profundidad en el espacio que el preparado por Shane-Wirtanen. Se basa en la información acumulada por Konrad Rudnicki y sus colegas, de la Universidad Jagelloniana de Cracovia. El mapa comprende galaxias de hasta una magnitud muy débil, a saber, las que representa la magnitud 20,5, con una distancia típica que viene a duplicar las de la muestra Shane-Wirtanen. Para esa distancia, hay tal cantidad de galaxias que resultaría inviable el estudio de un hemisferio celeste entero. Rudnicki y sus colegas concretaron su estudio a una región cuadrada del cielo, siendo de seis grados el tamaño de los lados del cuadrado. Incluso dentro de esta modesta superficie, llegaron a distinguir unas 10.000 galaxias. En el mapa de Rudnicki, los cúmulos no son los rasgos característicos. Muchos grupos y cúmulos de galaxias se superponen, de modo que en promedio desaparece el agrupamiento y la distribución de galaxias se aproxima a la distribución uniforme que se presupone en los modelos cosmológicos.

Una teoría de la distribución de galaxias debe guiarse por mediciones cuantitativas de la distribución. Se han utilizado dos enfoques generales para obtener tal información cuantitativa. Con cierta libertad podemos calificarlas como los métodos botánico y estadístico. En el



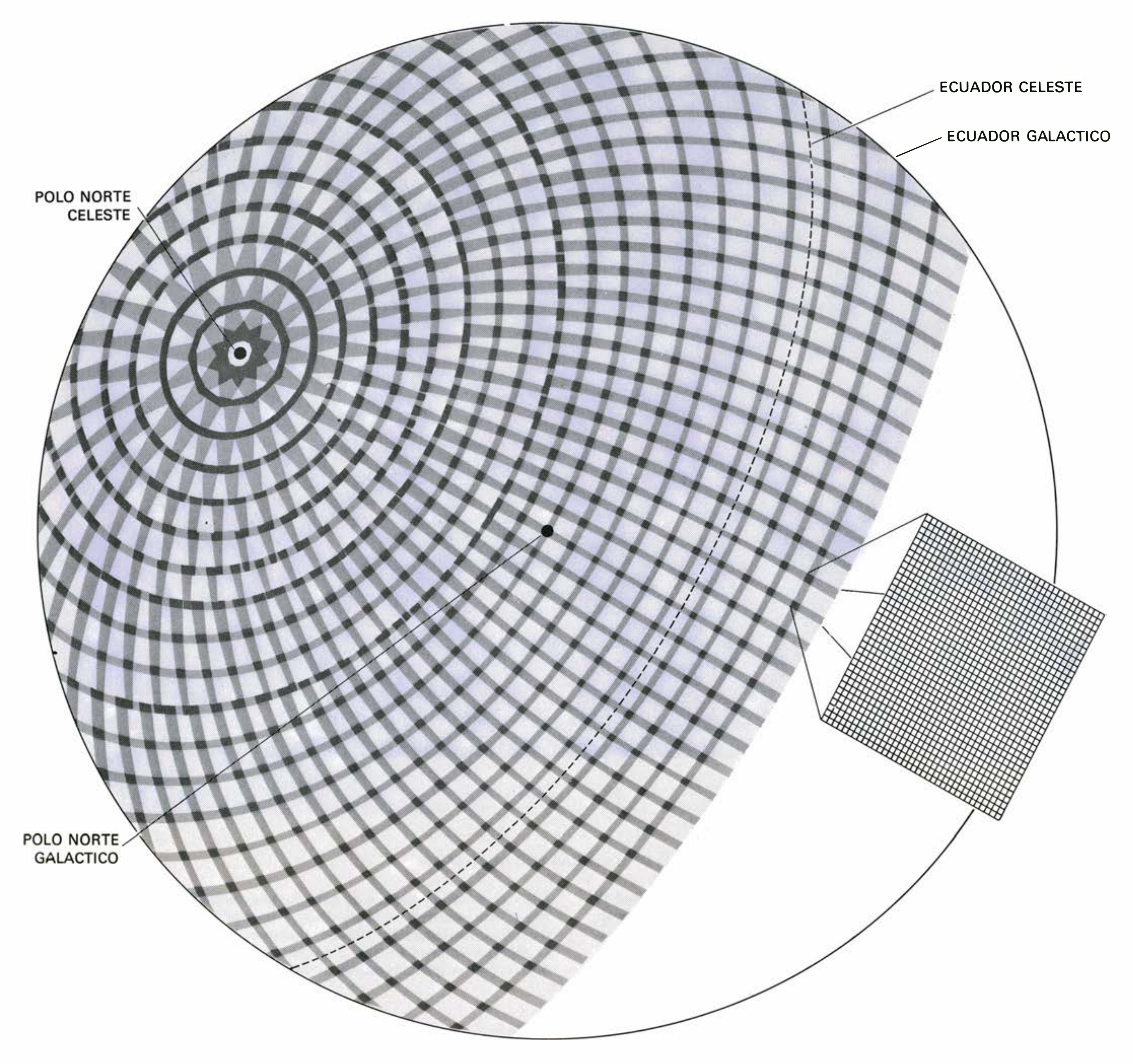
UN MILLON DE GALAXIAS con un brillo superior a la magnitud 19 se incluyen en un mapa que se extiende hasta una distancia promedio de 1400 millones de años-luz. El mapa se basa en las observaciones realizadas por C. Donald Shane y Carl A. Wirtanen; no recoge las posiciones de las distintas galaxias. Se ha construido dividiendo el cielo en cuadrados y contando el número de galaxias por cuadrado. Los números obtenidos se representaron mediante diversas intensidades de color gris. En el mapa sólo se distinguen los cúmulos mayores, debido a que muchos cúmulos se superponen unos a otros. Sin embargo, las propiedades estadísticas de la distribución de galaxias pueden medirse bien en el mapa de Shane-Wirtanen.

enfoque botánico, las agregaciones observadas de galaxias se clasifican de acuerdo con sus rasgos distinguibles. Por ejemplo, un aspecto destacado e importante consiste en los cúmulos muy grandes, como el cúmulo de Coma. George O. Abell, de la Universidad de California en los Angeles, catalogó casi 3000 cúmulos de esta clase, cubriendo una distancia aproximadamente igual a la de la muestra de Rudnicki. Abell halló que estos cúmulos estaban agrupados ellos mismos, a su vez; recientemente se ha descubierto que varios de los tales constituyen fuentes intensas de rayos X.

Aunque estas observaciones sean de gran interés, no parece que conduzcan a una teoría general de la agrupación de galaxias. La distribución global de galaxias es tan rica y variada que la clasificación de todos los grupos y agrupaciones de galaxias presenta un dudoso interés. En consecuencia, se ha adoptado el método estadístico, a modo alternativo.

Si queremos formular una teoría estadística necesitamos un medio para expresar en términos cuantitativos y simples la tendencia de las galaxias a formar cúmulos. Una hipótesis razonable con la que comenzar es suponer que las galaxias situadas en el interior de un cúmulo se encuentran más cercanas unas a otras que la distancia promedio. Esta tendencia se puede expresar cuantitativamente por medio de las funciones de correlación. La función de correlación a dos puntos mide la tendencia de un par de galaxias a hallarse más próximas que el promedio. Las funciones de correlación a 3 y 4 puntos miden la misma tendencia entre tríos y cuartetos de galaxias. En principio, se podría determinar la función de correlación para cualquier número de galaxias, pero no parece práctico ir más allá de la función a cuatro puntos.

Para calcular la función de correlación a dos puntos comenzamos por contar, de entre todas las parejas de galaxias recogidas en una muestra, aquellas que están separadas por una distancia determinada, por ejemplo, un grado. (En la práctica deben incluirse todos los pares que presenten una distancia comprendida dentro de un pequeño rango de distancias angulares, verbigracia, entre 0,8 grados y 1,2 grados.) Este número se divide luego por el número de tales parejas que se presentarían si las galaxias estuviesen distribuidas al azar, aunque uniformemente, por todo el cielo. Si la distribución fuera realmente uniforme, sin ningún tipo de agrupamiento, el resultado de la división sería 1; ya que se propone



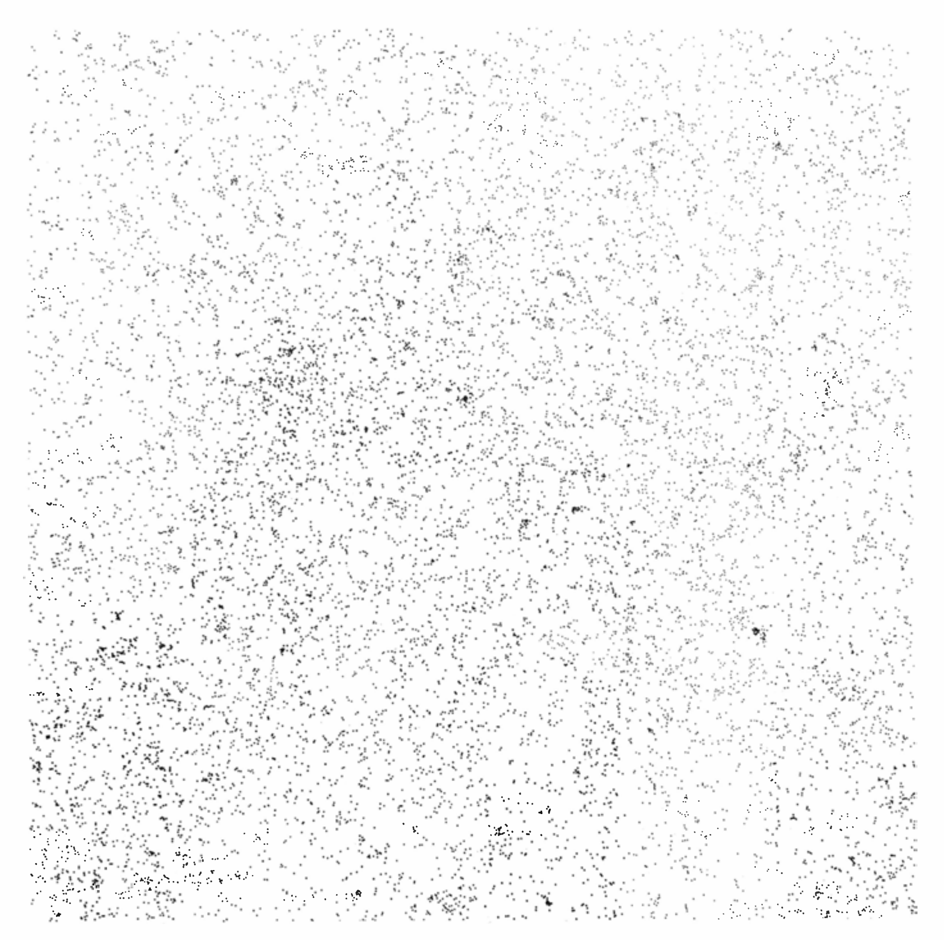
PROYECCION DEL CIELO utilizada en el mapa de Shane-Wirtanen; transforma las coordenadas de las galaxias desde un sistema de referencia basado en la orientación de la tierra a otro sistema de referencia basado en la orientación de nuestra galaxia. El mapa deriva de un estudio fotográfico en el que se dividió el cielo visible desde el Observatorio de Lick en secciones que se solapan, de seis grados por lado; cada sección corresponde a una placa fotográfica. Esta disposición en placas se centra en el polo norte celeste y se extiende hasta los 20 grados de Latitud Sur. Debido a que el eje de la tierra presenta una inclinación de 60 grados con respecto al eje de nuestra galaxia, todos los puntos del mapa deben ser desplazados en dicha cantidad, con el objeto de obtener un mapa basado en las coordenadas galácticas. La proyección resultante tiene como centro el polo norte galáctico y por circunferencia el ecuador galáctico. Ya que el perímetro del mapa corresponde al plano de la vía Láctea, en él se podrán observar muy pocas galaxias. Shane y Wirtanen dividen cada placa fotográfica en cuadrados de un sexto de grado por lado, contando entonces las galaxias que hay en cada uno de los cuadrados. El mapa se construyó con la ayuda de un ordenador.

esta condición para representar la no existencia de correlación entre pares, el resultado se ajusta a cero restando 1. El resto es la función de correlación a dos puntos para una separación angular de un grado; se trata de la razón de las parejas agrupadas a las parejas formadas accidentalmente por superposición. En ausencia de agrupamiento, la función es cero; si existen cúmulos la función es mayor que cero y mide la intensidad del agrupamiento.

En un análisis completo de la agrupación de galaxias, la función de correlación debe calcularse para un extenso rango de separaciones entre parejas. El mapa de magnitud 13, basado en los datos de Zwicky, no abarca un número suficiente de galaxias para una estimación fiable de la correlación; ahora bien, hemos calculado la función a dos puntos correspondiente a varios ángulos para la muestra de Zwicky de decimoquinta magnitud y para los mapas de Shane-Wirtanen y de Rudnicki. Muchos de nuestros hallazgos se basan en tales cálculos.

Una tendencia inequívoca es que la función de correlación disminuye conforme la separación angular que se estudia aumenta. En otras palabras, a mayor ángulo entre dos galaxias, menor probabilidad de que sea un par realmente agrupado. En la muestra de Zwicky, las galaxias separadas por 0,05 grados presentan a dos puntos una función de correlación de 10, aproximadamente. El significado de este número se puede interpretar de la siguiente manera: sobre 11 pares de galaxias con esa separación angular, 10 serán miembros de un cúmulo individual y no una agrupación accidental de galaxias con posiciones angulares similares, aunque diferentes distancias, a lo largo de la línea visual. Como la separación angular en la muestra de Zwicky se incrementa de 10 en 10 grados, cae rápidamente a 0,1; ello significa que sólo un caso sobre 11, con esta separación, pertenece a un cúmulo.

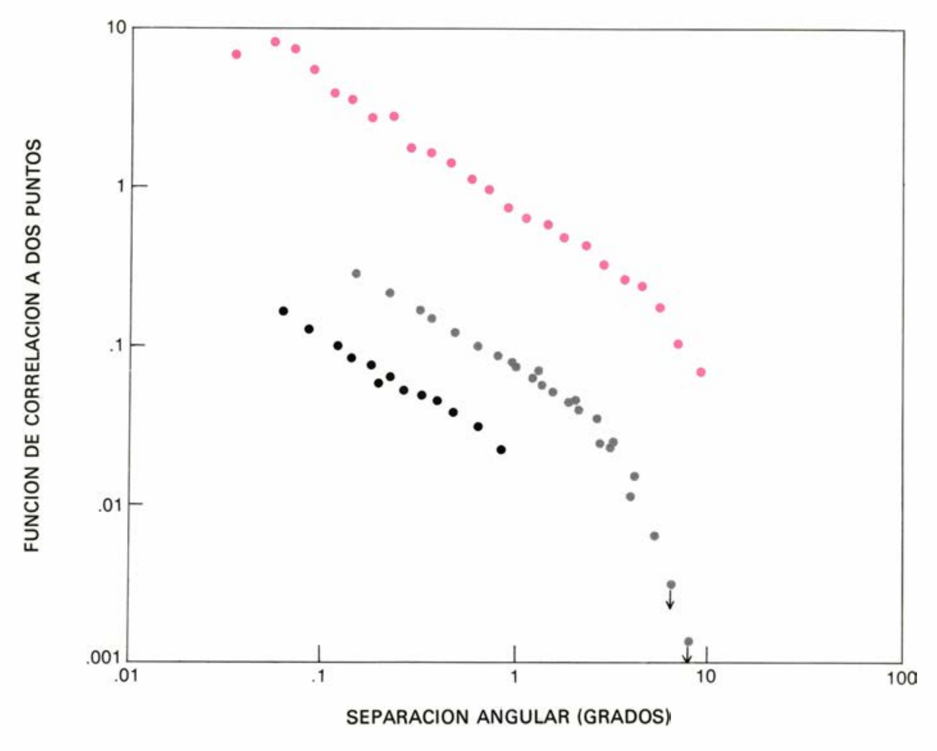
Una segunda tendencia manifiesta en los datos es que la intensidad de la agrupación parece disminuir conforme aumenta la profundidad de la muestra. Este resultado era de esperar, ya que la inclusión de galaxias más distantes incrementa necesariamente la superposición de cúmulos con distancias diferentes. Puesto que conocemos la distancia promedio de las galaxias de cada muestra, podemos calcular los efectos esperados como consecuencia del incremento de profundidad. Ya que un par típico de galaxias en la muestra de Shane-Wirtanen se halla 4,4 veces más alejado que un par equivalente en el mapa de Zwicky de magnitud 15, los pares de galaxias que presentan una separación angular de un grado en el mapa de Shane-Wirtanen deben compararse con los que, en el mapa de Zwicky, presentan una separación de 4,4 grados. De un modo análogo, galaxias que están separadas por un grado en el mapa de Rudnicki se comparan con las que se encuentran alejadas en 8,1 grados en el mapa de Zwicky. Mediante la aplicación de estos factores a las muestras más profundas podemos comparar las funciones de correlación de galaxias que posean la misma separación lineal.



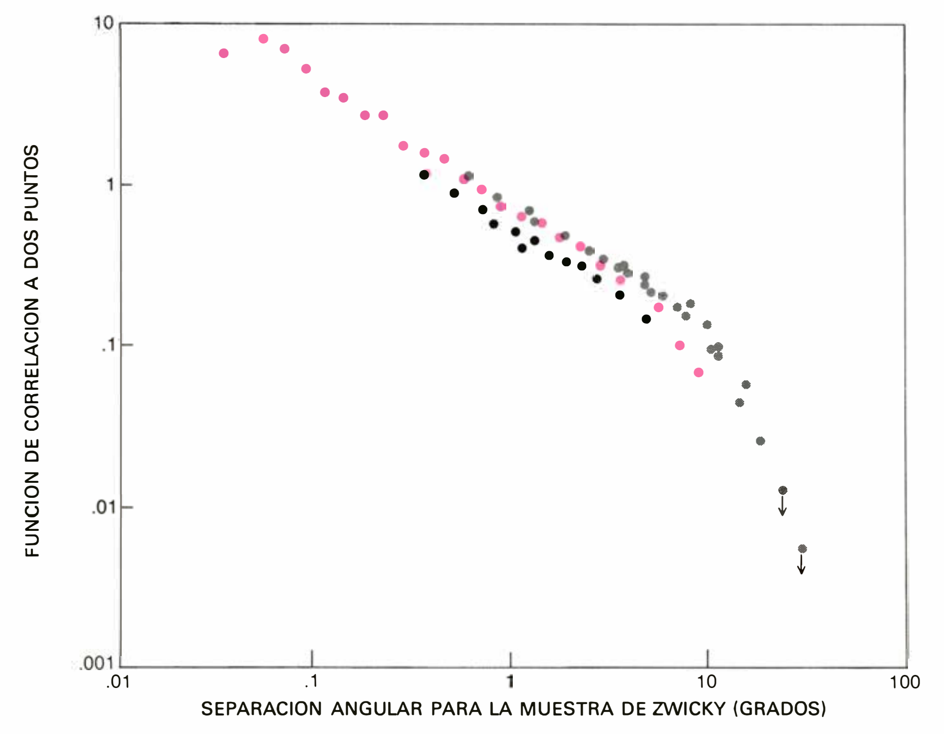
EL MAPA GALACTICO DE MAYOR PROFUNDIDAD, que comprende galaxias de magnitudes superiores a la magnitud 20,5, se funda en los datos obtenidos por Konrad Rudnicki y sus colegas de la Universidad Jagelloniana de Cracovia. A ese nivel de brillo hay demasiadas galaxias para que su distribución sobre el cielo pueda transformarse en un mapa, y la investigación se limita a una sección cuadrada de seis grados de lado; en superficie equivale a uno de los cuadrados grandes del diagrama de la página precedente. Incluso en una superficie tan pequeña hay más de 10.000 galaxias, con una distancia promedio a la tierra de 2800 millones de años-luz. El agrupamiento no es una característica señalada del mapa; por el contrario, la distribución se produce más bien al azar. La razón no consiste en que las galaxias más lejanas presenten una probabilidad menor de constituir agrupaciones, sino en el hecho de que coinciden más cúmulos en la misma línea de visión.

La intensidad con la que cúmulos independientes se solapan a lo largo de la línea de visión es también proporcional a la lejanía de la muestra. Un valor de la función de correlación de 1 en la muestra de Shane-Wirtanen es equivalente a un valor de 4,4 en el mapa de Zwicky; una función de correlación de 1 en la muestra de Rudnicki corresponde a un valor de 8,1 en el mapa de Zwicky. Así pues una medición realizada con la muestra de Shane-Wirtanen será equivalente a la hecha en la muestra de Zwicky cuando los ángulos medidos y las funciones de correlación calculadas se multipliquen por un factor de 4,4. Un ajuste similar, empleando un factor de 8,1, debe realizarse con la muestra de Rudnicki para poder relacionarla con las otras. (Se necesita realizar un pequeño ajuste adicional. No sólo vemos las galaxias en las muestras más profundas a mayores distancias, sino también a mayor precocidad temporal. Al hallarse el universo en expansión, los factores 4,4 y 8,1 deben modificarse ligeramente con el objeto de compensar el movimiento de las galaxias.)

Cuando se realizan estas correcciones, las funciones de correlación para las muestras de Shane-Wirtanen y de Zwicky coinciden casi exactamente. La correspondiente función para la muestra de Rudnicki presenta un conjunto similar de valores, pero ligeramente más bajos.



EL ALCANCE DEL AGRUPAMIENTO puede expresarse mediante las funciones de correlación, que miden la probabilidad de que un par o un grupo de galaxias sean miembros de un cúmulo dado. En esta figura se ha dibujado la función de correlación a dos puntos para la muestra de Zwicky de magnitud mínima 15 *(puntos coloreados)*, para la muestra de Shane-Wirtanen *(puntos grises)* y para la muestra de Rudnicki *(puntos negros)*. La función de correlación a dos puntos puede interpretarse como la razón de las parejas agrupadas, que están realmente próximas en el espacio una de la otra, con respecto a las parejas accidentales, que parecen juntas aunque se hallan a diferentes distancias a lo largo de la dirección de visión. Así, para la muestra de Zwicky sobre 11 parejas separadas por 0,05 grados, 10 están agrupadas, mientras que para 10 grados sólo una sobre 11 es una pareja agrupada.



LAS FUNCIONES DE CORRELACION AJUSTADAS reflejan las diferentes distancias promedio a las que se observan las galaxias en las tres muestras. Galaxias separadas por un grado en los tres mapas están mucho más alejadas (en medida lineal) en la muestra Shane-Wirtanen que lo que lo están en la muestra de Zwicky, y todavía lo están más en la muestra de Rudnicki. El número de cúmulos que se solapan a lo largo de la misma visual es también mayor para las muestras más profundas. Después de realizar el ajuste, debido a estos dos factores, las funciones de correlación medidas para las tres muestras presentan un buen acuerdo entre ellas. Conforme la separación entre galaxias aumenta, la función de correlación disminuye, y para ángulos pequeños se comporta como una línea recta. Según la pendiente de la recta, cada vez que la separación angular se dobla, la función de correlación disminuye en un factor fijo de 0,59. Más allá de 10 grados, la función de correlación disminuye con mayor rapidez, lo cual pone de manifiesto que, en el caso de distancias superiores de unos 60 millones de años-luz, el agrupamiento galáctico deviene relativamente débil.

Esta pequeña discrepancia no supone ningún problema; el estudio de Rudnicki es el menos exacto porque las distancias son tan grandes que se borra casi completamente la agrupación debido al emparejamiento accidental de las galaxias que se superponen. El acuerdo aproximado de las funciones de correlación para los tres mapas es un resultado importante y tranquilizador; indica que no hay serias perturbaciones en las funciones producidas por posibles errores en las observaciones. Más aún, proporciona una prueba cuantitativa de que el agrupamiento de galaxias muy distantes es estadísticamente idéntico al correspondiente a las galaxias próximas.

Una jerarquía de cúmulos

Para separaciones angulares pequeñas, la función de correlación a dos puntos presenta una dependencia sencilla con el ángulo. Si se representan en escalas logarítmicas los ángulos y las funciones de correlación, estas últimas disminuyen a lo largo de una línea recta conforme aumenta la separación angular. El hecho de que sea una línea recta indica que, cada vez que se dobla la distancia entre galaxias, la función de correlación disminuye en un factor fijo. La pendiente de dicha recta manifiesta que ese factor es 0,59.

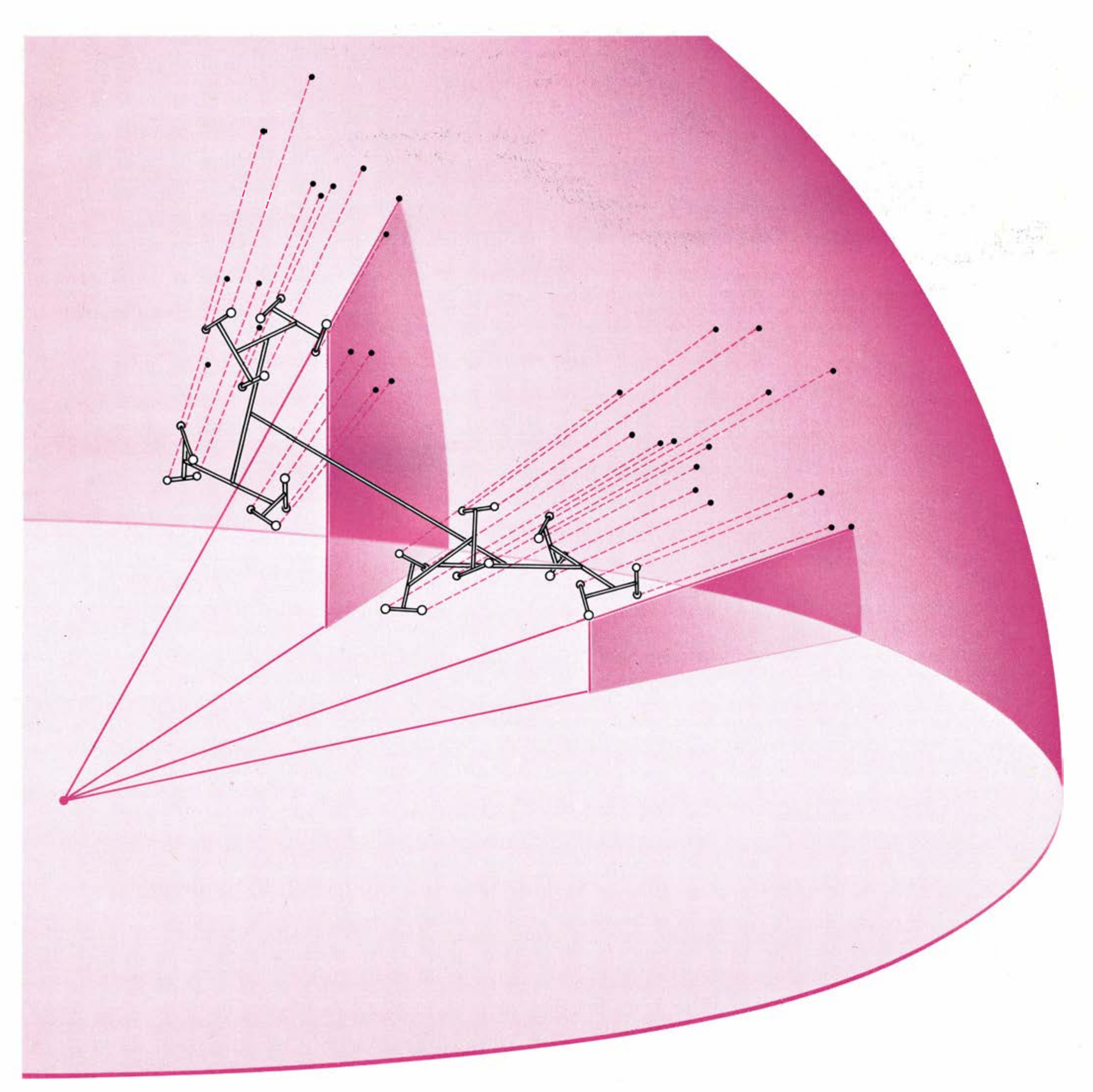
Para una separación de 10 grados en el mapa de Zwicky de magnitud 15 y para ángulos correspondientemente menores que para las otras muestras, la dependencia lineal de la función de correlación con la separación cambia bruscamente. Para estos ángulos tan grandes, la función de correlación disminuye más de prisa de lo que aumenta el ángulo. El punto para el que cambia la pendiente corresponde a una distancia lineal entre galaxias de unos 60 millones de años-luz. De esta manera, conforme aumenta la distancia entre galaxias, la probabilidad de que sean miembros de un cúmulo individual disminuye lenta y regularmente hasta que la separación es de 60 millones de años-luz; más allá de esta distancia, la tendencia de las galaxias a formar cúmulos decrece muy rápidamente.

Esta regularidad en la función de correlación fue el primer indicio de que la agrupación de galaxias debe seguir alguna pauta sencilla. Por otra parte, dicha pauta no queda especificada solamente a través de la función de correlación a dos puntos. Este hecho no resulta sorprendente, toda vez que la función de correlación a dos puntos condensa toda la información de los mapas galácticos en una sola gráfica.

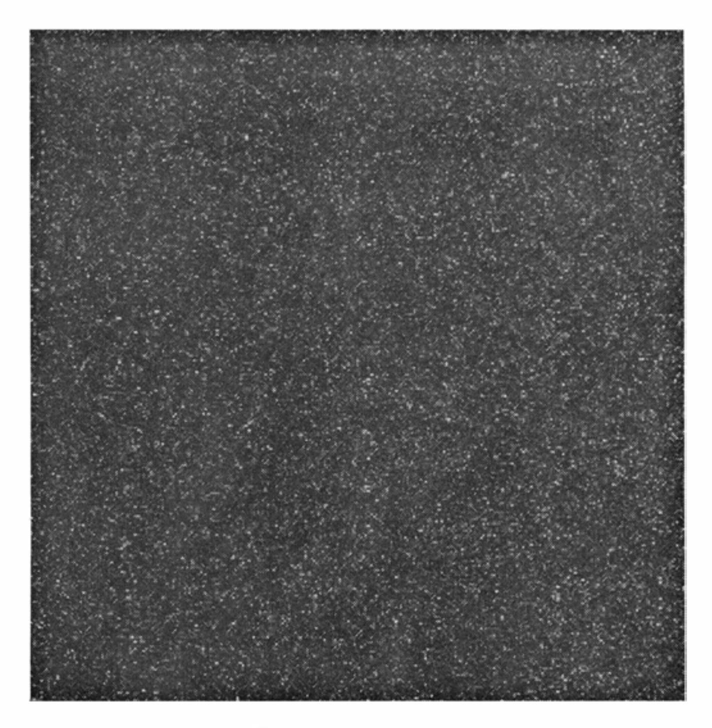
El rango de posibles interpretaciones

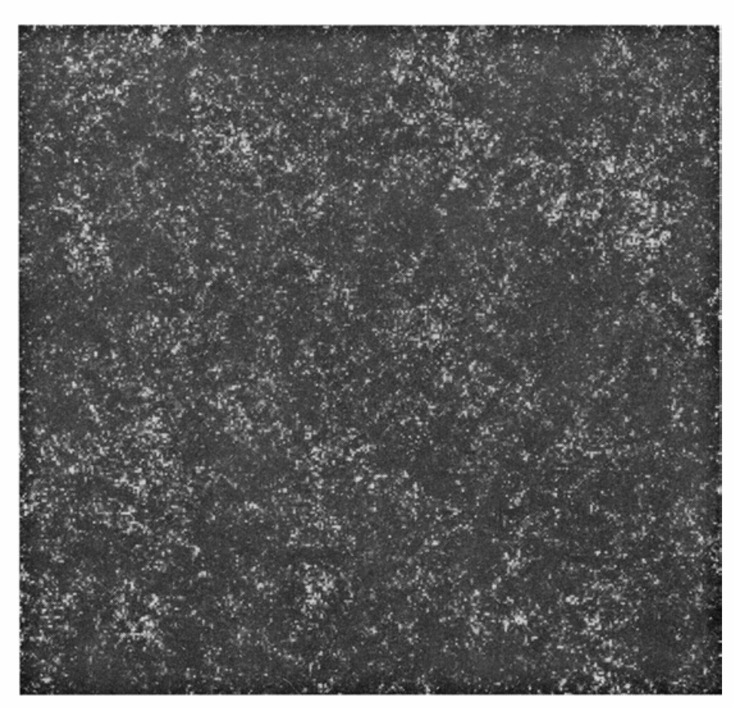
puede acortarse mediante la consideración de las funciones de correlación a tres y cuatro puntos. El cálculo de estas funciones se realiza contando el número de tríos y cuartetos de galaxias para cada una de las distintas separaciones angulares, y comparándolo después con el número de tríos y cuartetos que deberían esperarse para una distribución uniforme. Hemos descubierto que la función a tres puntos presenta también un comportamiento sencillo: es directamente proporcional al cuadrado de la función a dos puntos, y la constante de proporcionalidad resulta ser la misma para los tres enfoques. James Fry, uno de nuestros colegas en la Universidad de Princeton, halló que la función de correlación a cuatro puntos, en la muestra de Shane-Wirtanen, es proporcional al cubo de la función a dos puntos. (Las muestras de Zwicky y de Rudnicki son demasiado pequeñas para una estimación fiable de la función a cuatro puntos.) El significado físico de estos hallazgos puede establecerse de una forma sencilla: Si un buen lugar para hallar una galaxia es justo al lado de otra, un lugar mejor para mirar es cerca de un par de galaxias y, aún mejor, cerca de un trío de galaxias.

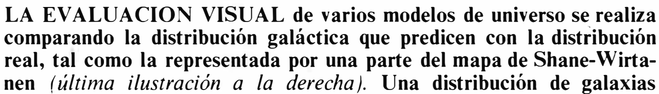
Teniendo presente esta información, podemos determinar el modelo de agrupamiento con una confianza razonable. Las galaxias se disponen en una jerarquía de cúmulos. Las galaxias suelen registrarse formando grupos pequeños y den-



MODELO DE UNIVERSO diseñado para reproducir las funciones de correlación medidas en los mapas de distribución de galaxias. El modelo se ha construido colocando una varilla imaginaria grande en una posición y orientación en el espacio aleatorias. En los extremos de ella, se emplazan pequeñas varillas orientadas nuevamente al azar; varillas todavía más pequeñas están ligadas a los extremos de estas últimas. En cada nivel, la longitud de las varillas se ha disminuido en un factor fijo elegido para reproducir las funciones de correlación a dos puntos observadas. Cuando se completa el número deseado de niveles se vinculan los extremos de las varillas más cortas con galaxias imaginarias. El resultado es una colección de galaxias distribuidas según una jerarquía de agrupación conocida. El modelo se construye con la ayuda de un ordenador, el cual, como paso final, proyecta las posiciones de las galaxias sobre el cielo de un observador imaginario. Se puede realizar entonces un mapa de su distribución. En la ilustración se muestra solamente un enjambre de puntos; un modelo que sea real exige varios miles de enjambres.







LA EVALUACION VISUAL de varios modelos de universo se realiza comparando la distribución galáctica que predicen con la distribución real, tal como la representada por una parte del mapa de Shane-Wirtanen *(última ilustración a la derecha).* **Una distribución de galaxias uniforme y completamente al azar, sin agrupación, proporciona una estructura** *(extremo izquierdo)* **marcadamente diferente de la estructura real; la diferencia de aspecto puede expresarse cuantitativamente por medio de la función de correlación, que para una estructura al azar**

sos; los grupos forman cúmulos; éstos, a su vez, integran cúmulos mayores y así sucesivamente hasta que el tamaño del cúmulo alcanza la escala para la que la función de correlación comienza a debilitarse intensamente, a unos 60 millones de años-luz. Más allá, el rango de agrupamiento es, sobre el promedio, comparativamente débil.

Aunque las pruebas cuantitativas que apoyan esta conclusión sólo se han obtenido recientemente, no es la primera vez que se ha propuesto en la astronomía una jerarquía de las estructuras. Ya en 1761, el matemático J. H. Lambert había sugerido la idea de la agrupación jerárquica. Fue recogida de nuevo por el escritor científico E. E. Fournier d'Albe en 1907 y, de nuevo en 1922, por el astrónomo C. V. L. Charlier.

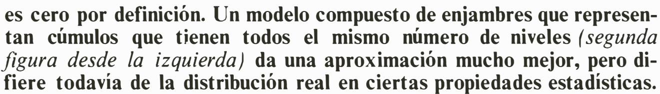
El descubrimiento en la década de 1920 de la verdadera naturaleza de las galaxias condujo a teorías más próximas a las nuestras. Un ejemplo es la observación de Abell de que cúmulos grandes, como el de Coma, tienden a formar cúmulos. Recientemente, el defensor más ferviente de la organización jerárquica fue Gérard de Vaucouleurs, de la Universidad de Texas en Austin. Hoy, nuestros métodos estadísticos han llegado a unas conclusiones parecidas. Sin embargo, media una importante diferencia entre la jerarquía de agrupación que se deduce de nuestros cálculos y las propuestas por investigadores anteriores. La jerarquía de agrupación, revelada por las funciones de correlación, no continúa indefinidamente, comienza a desaparecer cuando el tamaño de los cúmulos alcanza los 60 millones de años-luz.

Un modelo de universo

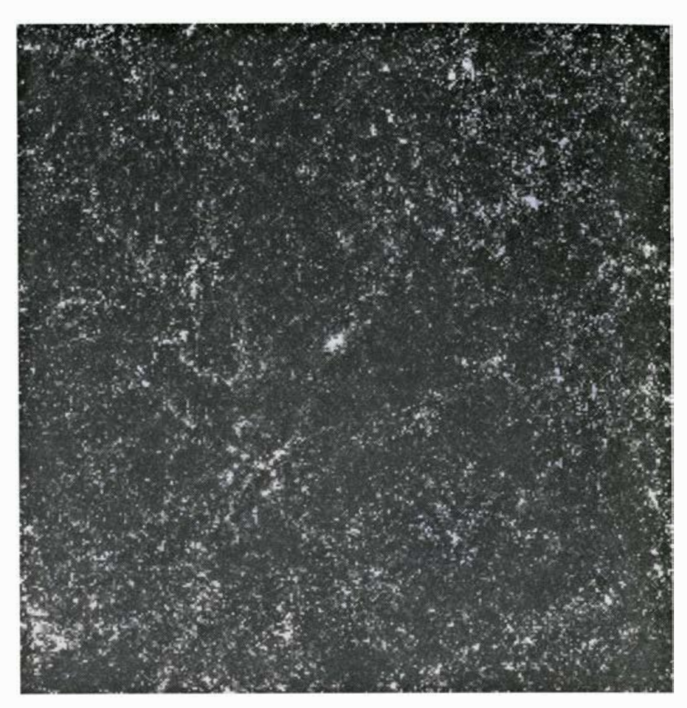
Si hemos interpretado correctamente las normas de agrupación galáctica, podremos construir un modelo de universo que presente los mismos rasgos que el universo real. Hemos construido tal modelo por un método que reproduce automáticamente las funciones de correlación observadas a dos puntos. En un punto escogido al azar, hemos colocado una varilla imaginaria orientada también aleatoriamente. En cada uno de sus extremos, colocamos varillas más pequeñas, orientadas asimismo al azar; la segunda serie de varillas presenta una longitud menor que la primera varilla según un factor fijo, que, de acuerdo con los cálculos, debe ser 0,57. En los extremos de las varillas de esta segunda serie hemos emplazado una tercera serie de varillas aún más cortas, con orientaciones asignadas de nuevo al azar y que son 0,57 veces más cortas que las del segundo nivel. Este procedimiento se continúa hasta que el tamaño de las varillas puede equipararse al de las galaxias; llegando a este punto se colocan, en los extremos de las varillas más cortas, galaxias imaginarias. Evidentemente, en la práctica no se construye esta especie de mecano de cúmulos de galaxias; en lugar de ello nosotros recurrimos a un ordenador para calcular las posiciones de las galaxias. Se colocan, en posiciones aleatorias, los cúmulos de galaxias a través de todo el modelo de universo; el ordenador se encarga de proyectar las galaxias sobre el cielo de un observador imaginario. El ordenador tiene en cuenta las variaciones en distancia y brillo intrínseco, por lo que muchas galaxias resultan demasiado débiles para ser vistas por el observador.

Este procedimiento resulta válido a la hora de reproducir el decrecimiento lineal observado en la función de correlación a dos puntos, debido a que la longitud de las varillas determina el número de parejas agrupadas que se pueden constituir para una determinada separación entre parejas. Se puede comprender por qué esto es así considerando en sentido inverso el orden jerárquico, es decir, desde las barras más pequeñas hasta las mayores. De una serie dada a la inmediatamente superior, la longitud de las varillas aumenta en el inverso de 0,57, es decir, en 1,75. Al mismo tiempo, se dobla el número de pares agrupados, ya que cada nivel superior conecta dos veces más galaxias mediante varillas rígidas. Por otra parte, el número de pares accidentales, no agrupados realmente, es proporcional al área del cielo estudiada; así pues, en cada nivel superior el número de tales parejas aumenta en un factor $(1{,}75)^2$ o 3,06. Puesto que la función de correlación es simplemente el cociente entre el número de pares agrupados y el





es cero por definición. Un modelo compuesto de enjambres que representan cúmulos que tienen todos el mismo número de niveles *(segunda figura desde la izquierda)* **da una aproximación mucho mejor, pero difiere todavía de la distribución real en ciertas propiedades estadísticas.**



Hay también diferencias apreciables para el ojo; en particular no hay suficientes regiones brillantes en el modelo. Esta deficiencia se corrige mediante un modelo que incluye enjambres-cúmulos de varios tamaños *(tercera figura por la izquierda).* **Este mapa semeja el mapa de Shane Wirtanen.**

de accidentales, la función de correlación disminuye en un factor 2/3,06, es decir, en 0,65, cada vez que la separación entre parejas aumenta en un factor de 1,75. Este ritmo es equivalente a un decrecimiento en un factor 0,59 cada vez que se dobla la separación entre parejas, lo cual coincide con el comportamiento que se observa en el universo real.

Este modelo de universo reproduce también las funciones de correlación a tres y cuatro puntos observadas, por lo menos de forma aproximada, ya que una galaxia que es miembro de n pares agrupados contribuye a su vez en $n^2/2$ tríos agrupados y $n^3/6$ cuartetos agrupados. (La aproximación es válida si n es grande.) Por tanto, las funciones de correlación a tres y cuatro puntos son proporcionales al cuadrado y al cubo de la función a dos puntos, tal como ocurre en el universo real.

En nuestro primer intento por construir un modelo de universo, llenamos el espacio imaginario con miles de enjambres de cúmulos estadísticamente idénticos. Cada uno de ellos poseía la varilla más larga de la misma longitud y un total de nueve niveles de varillas, con 512 galaxias en los extremos de las varillas más cortas. Cuando se proyectó sobre el cielo la distribución de galaxias realizamos un diagrama con el mismo formato gráfico que se empleó para la muestra de Shane-Wirtanen. El mapa resultante se distinguía fácilmente del de una distribución de galaxias al azar y presentaba algunas características interesantes, si bien atendiendo sólo a su apariencia resultaba decepcionante. El principal defecto de ese modelo estribaba en que existían muy pocas regiones brillantes, o, en otras palabras, había muy pocas agregaciones densas de galaxias. La diferencia podía detectarse en las funciones de correlación de órdenes superiores, ya que la contribución de los grupos densos en la cantidad de tríos y cuartetos de galaxias era mayor que en la de pares. Aunque las funciones de correlación a tres y cuatro puntos sí eran proporcionales al cuadrado y al cubo de la función a dos puntos, las constantes de proporcionalidad resultaban incorrectas: eran un factor 3 y 19, respectivamente, más pequeñas.

En un segundo modelo de universo, producimos más regiones densas mediante la inclusión de enjambres de cúmulos con diferente número de niveles y, en consecuencia, con un número de galaxias variante. Después de algunas pruebas concluimos con un modelo de 24.000 enjambres a siete niveles, 8000 de ocho niveles y cantidades menores de enjambres con más niveles, hasta los de doce niveles, de las que consideramos 40. Cada enjambre comienza con la varilla mayor del mismo tamaño, pero algunos cúmulos terminan antes que otros. Dentro de la precisión con que podíamos medirlas, las funciones de correlación a tres y a cuatro puntos resultaron idénticas a las de la muestra de Shane-Wirtanen. Más aún, la configuración de la distribución es similar al del mapa de Shane-Wirtanen. Pueden verse incluso algunos filamentos desgarrados como los que son característicos de los mapas del cielo. Muchos de nuestros colegas creen que estos filamentos son líneas reales o "pliegos" de galaxias en el espacio. No está claro en qué grado pueden ser comunes tales estructuras, pero el modelo ilustra el riesgo que suponen los juicios apresurados sobre estas materias. Los filamentos no se construyen en el modelo sino que resultan de accidentes en la ordenación aleatoria de las galaxias, acentuado por la sensibilidad que presenta la visión humana ante las características lineales. Algunos de los filamentos que se hallan en la distribución real de galaxias pueden ser artefactos de este tipo.

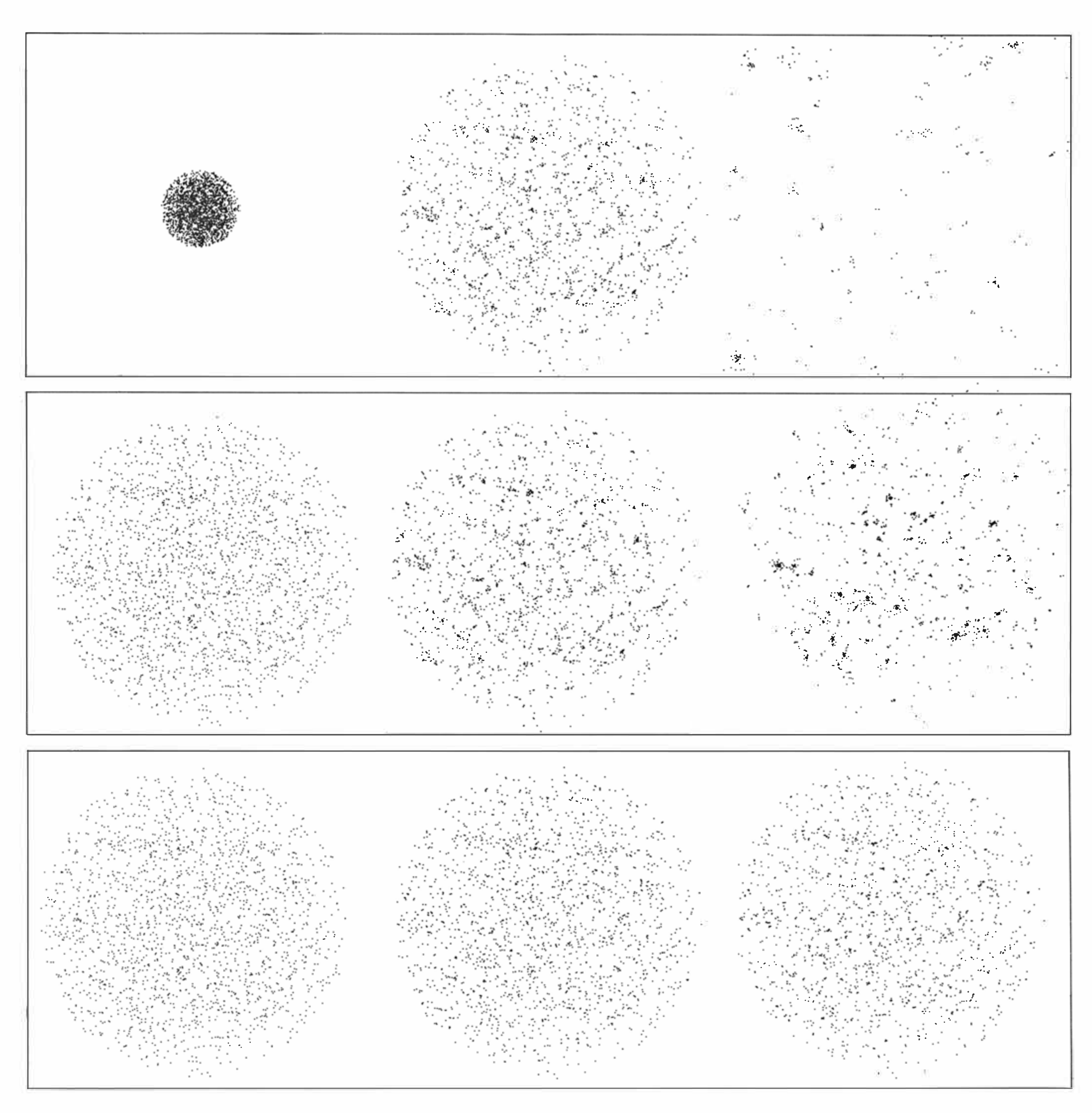
En nuestro modelo incorporamos intencionadamente una organización jerárquica, pero se ignoran algunos otros aspectos del agrupamiento. Por ejemplo, en una gran concentración de galaxias, como el cúmulo de Coma, las interacciones gravitacionales pueden ser tan fuertes que se rompan los pequeños grupos de galaxias, dando lugar a una distribución suavizada sin subcúmulos distinguibles. Sin embargo, sólo una pequeña fracción de todas las galaxias pertenecen como miembros a unos cúmulos de ese tamaño.

Un defecto más sutil de nuestro modelo es que el agrupamiento acaba bruscamente para una separación entre pares

algo mayor que la longitud de la varilla más grande, mientras que en el universo real persiste, para grandes escalas, un agrupamiento relativamente débil. Incluso con estos defectos, el modelo reproduce gran parte de la textura visual de la distribución de galaxias. El acuerdo entre el modelo y la observación sugiere que la consideración de una organización jerárquica de los cúmulos tiene que ser un enfoque razonable a la hora de describir la estructura a gran escala del universo.

La demostración de que hoy existe una jerarquía de los cúmulos de galaxias plantea inmediatamente la pregunta de cómo se ha ido estableciendo dicha jerarquía. Podemos describir una pauta posible de evolución que parece plausible a la luz de nuestros actuales conocimientos. Debe advertirse, sin embargo, que estas ideas pertenecen al terreno de la especulación teórica.

La base de la teoría de la "gran explo-



LA EVOLUCION DE CUMULOS a partir de universos homogéneos puede representarse mediante simulación con ordenadores. El modelo utiliza una muestra esférica del universo que contiene 2000 galaxias. La esfera de galaxias está en expansión de acuerdo con la teoría de "la gran explosión" sobre el origen del universo, pero dicha expansión se va debilitando por efectos gravitatorios. En la figura superior se ha representado la muestra poco después de "la gran explosión", después de expandirse según un factor 4 y después de expandirse en un factor 16. En la banda central se presenta la misma secuencia de sucesos pero con la escala ajustada artificialmente en cada caso, de modo que el tamaño de la esfera parece permanecer constante. Pequeñas fluctuaciones en la densidad son amplificadas por la expansión y dan lugar a cúmulos; con la subsiguiente expansión, los cúmulos, a su vez, se agregan entre sí para formar grandes cúmulos. La agrupación es el resultado de interacciones gravitacionales y, por tanto, está determinada por las masas de las galaxias, que en este modelo es lo suficientemente grande como para detener la expansión del universo. Masas más pequeñas producen una estructura significativamente diferente. En la secuencia evolutiva de la figura inferior la masa total es la mitad de la necesaria para detener la expansión, con el resultado de que se constituyen cúmulos pequeños pero no grandes cúmulos de cúmulos. La jerarquía de los cúmulos registrada de esta forma pone de manifiesto que el universo es más bien denso.

sión" para explicar la evolución cósmica estriba en la observación de que las galaxias remotas se alejan de nosotros con velocidades proporcionales a su distancia. En la teoría convencional de la gran explosión se puede hacer retroceder esta expansión general hasta un estado de la materia enormemente denso, que empezó a expansionarse hace unos 15.000 millones de años. La teoría requiere que la distribución original de la materia fuese casi exactamente uniforme: en el universo primitivo no existían cúmulos.

La evolución de los cúmulos

Una variable importante en esta teoría es la masa promedio de una galaxia, en otras palabras, la densidad media del universo. Si las galaxias son bastante masivas, entonces su mutua atracción gravitatoria puede llegar a detener la expansión, y el universo empezará a experimentar una contracción general. Si las galaxias no alcanzan cierta masa crítica, entonces la actual expansión del universo debe continuar indefinidamente. Como veremos, las masas de las galaxias tienen también una influencia decisiva en el desarrollo de los cúmulos de galaxias.

Hemos empleado un modelo computarizado en el estudio de la evolución del agrupamiento. El modelo comprende representaciones de 2000 galaxias que inicialmente están comprimidas a alta densidad en un volumen esférico del espacio. Las galaxias se distribuyen al azar, pero no uniformemente, y se les asignan velocidades que hacen que toda la muestra se expansione. Con la ayuda de un ordenador podemos calcular los movimientos posteriores de las galaxias bajo su mutua atracción gravitatoria. Una propiedad importante del universo real es que todo el espacio observado está ocupado por galaxias, pero nosotros no podemos reproducir estas propiedad en el modelo; nuestra muestra de galaxias se encuentra aislada. Al ignorar las galaxias exteriores, nuestra esfera seleccionada representa sólo una aproximación, pero es aceptable para cúmulos pequeños.

Conforme se va expansionando el modelo del universo y las galaxias se van alejando unas de otras, la gravedad frena la expansión. Las diferencias pequeñas y al azar en la densidad de galaxias se ven amplificadas: donde la densidad es ligeramente mayor que la densidad promedio, el campo gravitacional es ligeramente más intenso, y de esta forma la densidad disminuye más lentamente. Como consecuencia, se producen pequeñas irregularidades en la distribución original y nacen cúmulos. Cuando los cúmulos se forman, inicialmente su distribución es más o menos al azar, reflejando el hecho de que la distribución original era aleatoria. Sin embargo, los mismos cúmulos representan regiones con densidad mayor que la media, y, por tanto, la gravitación tiende a reunirlos en cúmulos algo mayores. Al repetirse el proceso a escalas más grandes, la expansión del universo crea una jerarquía natural de cúmulos.

El desarrollo de una jerarquía de agrupación en un modelo de evolución cósmica es un resultado alentador, pero es un resultado únicamente cualitativo. El modelo sería mucho más satisfactorio si pudiera realizar predicciones cuantitativas. Nuestro modelo, elaborado con ayuda del ordenador, no puede hacer tales predicciones debido principalmente a que es demasiado pequeño. Tratando la muestra de 2000 galaxias como si estuviera aislada, ignoramos gran parte del universo. Modelos mayores habrán de ofrecer una precisión mayor, pero, por ahora, no son viables. Cada una de las 2000 galaxias de nuestro modelo se mueve bajo la influencia de las fuerzas gravitacionales de las otras 1999; en consecuencia, para cada instante en el que se calculan las posiciones de las galaxias debemos evaluar 2000 × 1999 (casi cuatro millones) de interacciones. Si incrementamos el número de galaxias en un factor 10, el número de interacciones aumenta en un factor 100, lo que abrumaría al más rápido ordenador.

Un gas de galaxias

Debido a que una correcta simulación de la evolución cósmica parece estar más allá de nuestros medios, nos hemos enfrentado con el problema desde otro enfoque, en el que las galaxias de un universo en expansión son consideradas como las partículas de un gas. Esto representa también una aproximación del problema real, pero puede ser más fructífera. En la física del plasma (disciplina que se ocupa del comportamiento de gases a alta temperatura e ionizados) se ha desarrollado una teoría que describe la agregación de partículas en un gas. Paradójicamente, dicha teoría no ha supuesto ningún rendimiento para la física de plasmas; en su forma exacta, se hace demasiado complicada para los resultados que con ella pueden calcularse, pero las versiones simplificadas se muestran inadecuadas para describir muchos fenómenos de interés. Aplicando dicha teoría a la evolución de cúmulos de galaxias tenemos una importante ventaja: ya conocemos ciertas propiedades importantes del "gas", tales como las funciones de correlación a dos, tres y cuatro puntos. Con esta base podemos intentar simplificar la teoría sin descartar sus detalles esenciales. Marc Davis, de la Universidad de Harvard, que trabaja en colaboración con nosotros, ha completado recientemente un estudio de evolución cósmica empleando este método.

Los resultados obtenidos concuerdan con dos características esenciales de la distribución observada de las galaxias: la función de correlación a tres puntos predicha tiene la magnitud correcta con respecto a la función a dos puntos, y esta última sigue, logarítmicamente, una línea recta para decaer después de forma abrupta.

La coherencia de estos resultados apoya nuestro modelo del desarrollo de la jerarquía de agrupación observada. No todas las observaciones, sin embargo, se manifiestan coherentes con este modelo; hay una medición de gran importancia que contradice directamente nuestros resultados.

Como se ha señalado antes, la masa media de las galaxias tiene una fuerte influencia sobre el desarrollo del agrupamiento. En los cálculos basados en la teoría prestada por la física de plasmas, la masa obtenida fue justamente la necesaria para detener la expansión. Sin embargo, estimaciones de las verdaderas masas de las galaxias sugieren que son demasiado ligeras para detener la expansión en un factor comprendido entre 10 y 30.

Ahora bien, si tomamos en nuestro modelo masas tan pequeñas obtenemos predicciones un tanto distintas: se formarán pequeños y densos cúmulos, pero no grandes cúmulos de cúmulos. La jerarquización cesa mucho antes, o lo que es lo mismo, las funciones de correlación disminuyen más rápidamente conforme aumenta la separación.

Por el momento, no podemos dar cuenta de la discrepancia entre las masas necesarias, al parecer, para explicar la jerarquía de los cúmulos de galaxias y las masas que los astrónomos han estimado por métodos independientes. Por otra parte, no nos desalienta dicha contradicción, porque, en ambos cálculos, hay muchas incertidumbres. El conflicto resulta de lo más intrigante. Si las galaxias tienden a ser más masivas, tenemos formulada una teoría satisfactoria de la agrupación de las galaxias, pero permanece pendiente el problema de localizar la fracción de masa que falta. Por el contrario, si las masas estimadas son correctas, nos enfrentamos con otro problema, igualmente interesante: hallar otra vía que explique la pauta sencilla que subyace en el agrupamiento de las galaxias.

Ciencia y sociedad

Nobel de Física 1977

La primera reacción que expresamos quienes nos ocupamos del campo de la física del estado sólido, al escuchar los nombres de los científicos galardonados con el premio Nobel de Física, fue de satisfacción y total acuerdo. La segunda, de recriminación por la injusta tardanza. La lista de contribuciones de primera magnitud de los profesores Van Vleck, Mott y Anderson es tan extensa que ya no era cuestión de sí o no, sino de cuándo recibirían el premio Nobel. Anderson, a sus 53 años, es el joven nexo de unión de dos grandes veteranos de la física de nuestro siglo: Van Vleck, de 78 años, y Mott, de 72. Fue alumno del primero en Harvard y ha colaborado intensamente con Mott en el laboratorio Cavendish durante estos últimos años. El premio Nobel de este año se ha otorgado a toda una vida científica; a tres hombres que durante años han estado, y ¡están! abriendo nuevos campos, trabajando en la primera línea de los ya abiertos y dirigiendo con sus ideas gran parte de la física del estado sólido.

John Van Vleck se doctoró, en 1922, en la Universidad de Harvard, a la edad de 23 años, volviendo a ella como profesor en 1934. Permanecerá asociado a Harvard toda su vida, con intervalos de ausencia como profesor invitado en numerosas universidades del mundo. Inició su carrera como investigador trabajando en espectroscopía óptica y relaciones de dispersión dentro del formalismo de la vieja teoría cuántica, temas sobre los que publicó su primer libro: "Quantum Principles and Line spectra".

A partir de ese momento, Van Vleck se dedica al estudio de susceptibilidades eléctricas y magnéticas. Su libro "Theory of Electric and Magnetic Susceptibilities" es la primera aplicación rigurosa de la mecánica cuántica al cálculo de efectos magnéticos en sólidos. Esta obra marca un hito en la historia del magnetismo; se trata de un caso extraordinario en la historia de la física; escrito en 1932, su lectura resulta, todavía, fascinante y actual. Dos son, a mi entender, las razones de este hecho insólito: por un lado, el acierto del autor que supo escribirlo en un momento en que el formalismo teórico estaba suficientemente bien establecido y, por otro, la especial habilidad de Van Vleck, no sólo para expresar sus propias ideas, sino también para iluminar y dar forma física sencilla a ideas generadas por otros investigadores: la teoría de Félix Bloch de ondas de espín, el concepto de campo cristalino introducido por Hans A. Bethe, el llamado Hamiltoniano de canje de Heisenberg adquieren una forma más clara y físicamente visualizado en sus manos. Van Vleck es realmente el responsable del Hamiltoniano de la forma $J \cdot \Sigma(Si \cdot Sj)$ que hoy usamos para describir aislantes magnéticos, aunque fuesen Werner Heisenberg y P.A.M. Dirac los primeros en darse cuenta de la conexión entre canje y ferromagnetismo y entre canje y el producto escalar de operadores de espín, respectivamente.

Se debe a Van Vleck una expresión general de la susceptibilidad paramagnética independiente de la temperatura aplicable a iones, átomos y moléculas: el llamado paramagnetismo de Van Vleck. Las medidas experimentales realizadas en el laboratorio del físico español Blas Cabrera fueron la primera confirmación experimental de la validez de la teoría de Van Vleck. La explicación del fenómeno de anisotropía magnetocristalina, basada en la ligadura de los espines a la red mediante el acoplamiento espín-órbita, es otro de los logros de Van Vleck, autor asimismo de la primera teoría destinada a explicar el fenómeno del supercanje. Otra parte importante del trabajo de Van Vleck se refiere al llamado efecto Jahn-Teller, asociado a la interacción entre los electrones y la posición y movimiento de los núcleos atómicos. Algunas veces un átomo extraño puede reemplazar un átomo del sólido sin que ocurran cambios sustanciales en la estructura de la red cristalina. Bajo determinadas circunstancias, la estructura electrónica del átomo extraño es tan incompatible con la simetría de los átomos circundantes que produce una distorsión local de la red. Este efecto, llamado de Jahn-Teller, fue predicho en 1930, pero sólo durante la última década, y debido esencialmente al trabajo de Van Vleck, se ha llegado a entender en detalle este fenómeno y su importancia experimental.

La colaboración de Van Vleck y el grupo holandés de bajas temperaturas fue crucial para el entendimiento de fenómenos de relajación de vectores magnéticos, espín-espín y espín-red, siendo el propio Van Vleck la figura central en la aplicación de estos conceptos a la espectroscopía de radiofrecuencias, después de la segunda guerra mundial.

A la vista de tales contribuciones, no es de extrañar que pueda considerársele como "el padre del magnetismo moderno". De hecho, y hasta hace unos años, magnetismo y Van Vleck eran prácticamente sinónimos. Hoy en día, la física del magnetismo se ha expandido tanto que nadie puede abarcar todas sus ramas de la forma en que él lo hacía en sus comienzos. Con palabras de Phil Anderson, pronunciadas en la sesión en honor de Van Vleck, durante la conferencia de Boston sobre magnetismo en 1967: "Un hombre, John Van Vleck, mediante su trabajo en la teoría del campo cristalino, paramagnetismo, resonancia espectroscópica y teoría cuántica ha convertido el magnetismo en una ciencia demasiado amplia para cualquier otro hombre".

Sir Nevill Mott termina su B.A. en Cambridge en 1926, el mismo año en que aparece la ecuación de Schrödinger. El laboratorio Cavendish de Cambridge será el eje de su vida científica durante sus primeros años. A lo largo de los años 1933-1953 enseña física en la Universidad de Bristol. En 1953, vuelve a Cambridge como "Cavendish Professor", puesto que ocupa hasta su jubilación en 1971. Todos sus antecesores en esta cátedra, famosa mundialmente, han recibido el Premio Nobel, excepto, claro está, el primer ocupante J. C. Maxwell. En la actualidad es profesor emérito de física en el Cavendish. Mott ha permanecido durante toda su vida profundamente interesado por los problemas de la educación, en todos los niveles.

Resulta realmente difícil hacer justicia en pocas líneas a los logros científicos de un hombre que ha venido publicando durante 50 años en la vanguardia de

tantos campos y que en la actualidad, a sus 72 años, continúa trabajando con la misma ilusión y productividad de siempre, sin mostrar el mínimo signo de cansancio o anquilosamiento. Sus primeras contribuciones importantes tienen que ver con la teoría del scattering atómico (colisiones atómicas), incluyendo el darse cuenta del papel de la simetría en la colisión de partículas α con núcleos de He. La teoría de Mott permitió realizar predicciones muy concretas, que fueron verificadas experimentalmente por el descubridor del neutrón James Chadwick. En 1934, publicó, en colaboración con Massey, su libro "The Theory of Atomic Collisions". Esta obra es el primer ejemplo de lo que llegó a ser una de las características de Mott: escribir un texto clásico, una referencia obligada hasta nuestros días, antes de saltar a un nuevo tema, en este caso, los metales.

En 1933, Mott se trasladó de Cambridge a Bristol; allí, bajo la influencia de Herbert Skinner y Harry Jones, empezó a trabajar en lo que ha sido su campo de investigación desde entonces: la física del estado sólido. La etapa Mott-Jones terminó con la publicación de su libro conjunto "The Theory of the properties of metals and alloys"; a ésta siguió su colaboración con Ronald Gurney, con el que desarrolló su teoría de procesos fotolíticos en emulsiones fotográficas. Durante los años de su colaboración con Gurney, su interés se centró principalmente en el estudio de las propiedades de los materiales aislantes. Su libro, en colaboración con Gurney, "Electronic processes in ionic crystals" es, una vez más, un clásico en la física de los aislantes. Una mirada a la lista de sus publicaciones después de la guerra nos muestra a un Mott desarrollando nuevas ideas sobre problemas tan complejos y difíciles como son los defectos y dislocaciones en sólidos. En 1948 establece, en colaboración con Nicolás Cabrera, la famosa teoría Cabrera-Mott sobre oxidación superficial.

Hacia la mitad de los años cincuenta, Mott comenzó a interesarse en varios campos nuevos: metales líquidos, impurezas en semiconductores, materiales desordenados y semiconductores amorfos. Pronto, durante esta fase, resuelve un problema que había permanecido sin solución durante largo tiempo: la transición de comportamiento aislante a metálico al crecer la densidad de portadores de carga; es decir, la llamada transición de Mott. Las propiedades eléctricas de los cristales se describen mediante la teoría de bandas que los clasifica, de



Phil Anderson, de 53 años, es el joven nexo de unión entre Van Vleck y Mott

acuerdo con su conductividad eléctrica, en metales, semiconductores y aislantes. Sin embargo, esta teoría no es válida universalmente; una llamativa excepción es el óxido de níquel que es aislante, mientras que según la teoría de bandas debería ser un conductor tipo metálico. Mott ha demostrado la viabilidad de que esto ocurra mediante una teoría más refinada que toma en cuenta la interacción entre los electrones. Para entender la idea física de este fenómeno, imaginémonos una serie de átomos de hidrógeno; quitemos un electrón a uno de ellos; el protón resultante atraerá a dicho electrón y podía obligarle a "situarse" en su estado ligado, siendo $a_0=0{,}529$ angstroms ($1\text{Å}=10^{-8}$ centímetros) la separación entre el protón y el electrón. Supongamos ahora que ya hay presentes otros electrones semejantes. La presencia de dichos electrones libres altera la descripción anterior acerca de la ligadura del electrón al protón; dichos electrones apantallarán el potencial de Coulomb visto por nuestro electrón, reduciéndolo. Si la longitud de apantallamiento (dependiente de la densidad de electrones libres) es menor que a_0, el ion no puede atrapar al electrón, que queda libre con una función de onda que se extiende por todo el sólido y, por tanto, es capaz de conducción eléctrica. Este electrón contribuirá a su vez a que el apantallamiento sea más efectivo, colaborando en la liberación de otro electrón; éste ayudará a liberar a otro, y otro a otro, y así sucesivamente. La transición una vez empezada será muy rápida.

Citemos, por nombrar una más de sus muchas contribuciones durante este período, el llamado excitón de Mott. Un excitón puede visualizarse como un estado ligado de un electrón en el potencial de un hueco. En el excitón de Mott, la función de onda del electrón se extiende sobre varios átomos (a diferencia del excitón de Frenkel, en el que existe una probabilidad elevada de encontrar el electrón y el hueco en el mismo átomo del cristal). Se han hallado excitones de Mott en cristales semiconductores con intervalos de energía prohibidos ("gaps") pequeños y constantes dieléctricas grandes.

En 1958, Mott captó inmediatamente la importancia del famoso artículo de Anderson sobre ausencia de difusión en ciertas redes desordenadas, y desde entonces concentró sus esfuerzos en el estudio de la física de sistemas desordenados. Mott generalizó la idea original de

Anderson mostrando la posibilidad de existencia de estados extendidos y localizados en el mismo cristal dependiendo de la energía del estado. Su libro más reciente, escrito con Davis, "Electronic processes in non-crystalline materials" basta para mostrar la impresionante magnitud de sus logros, así como, dicho por Mott, "el amplio camino que queda por recorrer para alcanzar un buen entendimiento de la física de los sólidos desordenados"

Quizá la característica más resaltable de Mott sea su profundo interés por los resultados experimentales, por problemas físicos reales, cualesquiera que sea su complejidad y dificultad, huyendo siempre de modelos teóricos ficticios. Esta característica le convierte en especialmente atrayente para los físicos experimentales, a pesar de la evidente dificultad que encuentran para seguir su línea intuitiva de razonamiento (dificultad que, por otro lado, no es menor para sus colegas teóricos).

Su estilo intuitivo, tan peculiar, de hacer física le permite avanzar en problemas donde mentes quizá tan potentes, pero más convencionales, se estancan.

Philip Warren Anderson, o Phil Anderson como él prefiere ser llamado, se doctoró (bajo la dirección de Van Vleck) en Harvard en 1949, y ese mismo año empezó su colaboración con los Bell Telephone Laboratories, colaboración que todavía mantiene. Desde 1967 hasta 1975, dividió su tiempo entre el Cavendish Laboratory de la Universidad de Cambridge y los Bell Telephone Labs. En la actualidad, su trabajo se reparte entre los Bell y la Universidad de Princeton.

Después de terminar su doctorado en Harvard, su interés se centró en varios aspectos del magnetismo. Su artículo, escrito en 1959, "New Approach to the theory of Superexchange Interactions" presenta una teoría de supercanje más general y completa que la de Van Vleck. La idea clave es la interacción de los electrones *3d* con los átomos diamagnéticos.

En otro artículo, de gran originalidad e importancia, publicado en 1961, Anderson explicó por qué los átomos de hierro son magnéticos cuando se disuelven en un material no magnético, como el cobre, por ejemplo. Este importante descubrimiento (es decir, el de los momentos magnéticos localizados o "miniimanes") ha allanado el camino para el entendimiento, a nivel microscópico, del origen del magnetismo en los metales de transición tales como el hierro, níquel, cobalto y manganeso, así como para entender por qué otros metales como el cobre y la plata no son magnéticos.

El magnetismo del hierro es un factor esencial en nuestra civilización moderna. Sin hierro no habría generadores, motores y transformadores eléctricos necesarios para la transformación de energía eléctrica a largas distancias.

Anderson publicó, en 1958, su famoso artículo sobre la ausencia de difusión en ciertas redes desordenadas. Estudia allí el movimiento de un electrón en un cristal, en el que la probabilidad de salto desde una posición hasta otra contigua es la misma para todos los puntos de la red, pero en el que la energía del electrón varía de una posición a otra en la red. Para una variación pequeña, el electrón puede desplazarse libremente a través del sistema, dando lugar a funciones de onda que se extienden a lo largo del cristal y, por tanto, es capaz de conducción eléctrica. Sin embargo, cuando las fluctuaciones de la energía del electrón en los puntos de la red son grandes, hay un nuevo tipo de transición hacia un estado aislante: los electrones quedan atrapados con funciones de onda localizadas en la vecindad de un determinado átomo. El modelo de Anderson, originalmente diseñado para explicar resultados experimentales sobre difusión de espines no nucleares, es en la actualidad de crucial importancia en la física de semiconductores amorfos, donde el desorden en la estructura se traduce en que los electrones son atraídos más fuertemente a unos átomos de la red que a otros. Con palabras de Sir Nevill Mott, pronunciadas en 1976: "El artículo de Anderson, publicado en 1958, a pesar de haber sido entendido por pocos y no ser creído por muchos se ha convertido, una década más tarde, en la explicación de casi todas las propiedades eléctricas de los materiales amorfos".

Hacia 1958, Anderson empezó a interesarse también por el fenómeno de la superconductividad y, a través de ella, por la teoría de muchos cuerpos. La teoría B.C.S. (por la que Bardeen, Cooper y Schrieffer recibieron el premio Nobel en 1972) adquiere una forma sencilla y elegante mediante la formulación de Anderson en términos de pseudo-espines. En dicha formulación ya está inherente el concepto de simetría rota ("broken symmetry") en superconductores, idea que Anderson desarrolla y explica en su libro "Concepts in solids", de obligada lectura para todo aquel interesado en las ideas físicas fundamentales sobre las que se basa la ciencia del estado sólido. En 1962, explica lo que entiende por simetría rota en su primer curso para graduados en el Cavendish; entre sus estudiantes se encuentra Brian D. Josephson quien, son sus propias palabras, "fascinado por el concepto de simetría rota, me planteé una pregunta obvia: ¿Cuál es el efecto de la simetría rota en superconductores? ¿Hay alguna forma de detectarla experimentalmente?". La respuesta que Josephson encuentra a esta pregunta, y que animado por Anderson publica en 1962, constituye la esencia del hoy llamado efecto Josephson, por el que obtendría el premio Nobel de Física en 1973 (compartido con Leo Esaki e Ivar Giaever).

Las contribuciones de Anderson a la ciencia de semiconductores, líquidos cuánticos (tales como ^3He y ^4He), el llamado efecto kondo, "vidrios de espín" y estrellas de neutrones, por citar algunos, son de vital importancia. De hecho, y como muchos físicos ironizan: tema en el que Anderson se interesa, tema en el que deja su marca, su contribución genial.

A título anecdótico, y puesto que el título de esta sección es "Ciencia y sociedad", es interesante señalar que fue Anderson quien inició en Cambridge un curso especial titulado: "Ciencia, Tecnología y Sociedad"; dicho curso todavía continúa, tras vencer la primera oposición por parte de muchos incrédulos. El programa inicial de Anderson convence a todos de que los problemas planteados por la ciencia y la tecnología a nuestra sociedad y viceversa son muy dignos de su inclusión en una carrera de Ciencias al mismo nivel que la óptica, mecánica o electricidad, por ejemplo.

Phil explicaba, con su estilo directo y característico, que un científico es esencialmente alguien que intercambia información verificable por reconocimiento y fama. Esta imagen, que a algunos nos parecía excesivamente pesimista, se ha convertido, con la experiencia, en triste realidad, al menos estadísticamente. Las charlas informales de Anderson con Heine, Mott y todos los que querían tomar con él una taza de café los sábados en el Cavendish, son recuerdos imborrables tanto de un hombre dispuesto a aconsejar al estudiante ávido como de un genio en acción.

A modo de resumen, haremos hincapié en esa característica común de estos tres físicos teóricos: su contacto constante con los experimentos, su avidez por conocer los nuevos resultados experimentales. Son teóricos que saben hablar el lenguaje de los experimentales; prueban, con su propio ejemplo, la veracidad de la afirmación de Mott: "los

laboratorios en los que teóricos y experimentales aprenden a hablar uno el lenguaje del otro, son los que progresan más rápidamente". (P.M.E.)

Nobel de Medicina 1977

El premio Nobel de Medicina correspondiente al año pasado se ha concedido a los profesores Guillemin, Schally y Yalow, bioquímicos investigadores en la naturaleza de las hormonas.

El conocimiento de las relaciones entre cerebro y hormonas se remonta a varios decenios atrás. Aunque en los comienzos de la Endocrinología, en los tiempos de Brown-Séquard, parecía que las glándulas de secreción interna eran un mosaico de piezas sueltas, pronto se supuso un control superior y la existencia de un sistema endocrino: se identificaron las trofinas, tropinas o estimulinas de origen pituitario; asimismo, se estableció la función rectora central que correspondía a la hipófisis en el sistema.

Pero todavía en 1926, el fisiólogo Gley afirmaba que el funcionamiento de las glándulas endocrinas era independiente del sistema nervioso. Posteriormente, la observación de los reflejos opto-pigmentarios y opto-sexuales, así como los estudios anatómicos, pusieron de relieve la íntima relación que existe entre el sistema endocrino y el sistema nervioso central; luego se entrevió, y en seguida confirmó, la estrecha vinculación, no sólo anatómica, de la hipófisis con el diencéfalo o hipotálamo, situado en la base del cerebro y sede también de las emociones y de los núcleos vegetativos centrales en el sistema nervioso autónomo. El diencéfalo pasó a ser el engarce y la encrucijada en la que el psiquismo enlazaba con las funciones viscerales a través del sistema nervioso vegetativo y del sistema endocrino. Desde entonces en adelante se habló del eje diencéfalo-hipofisario como del enclave neuro-endocrino superior.

Durante mucho tiempo, sin embargo, se tuvo al lóbulo posterior de la hipófisis como el único relacionado con el diencéfalo. Los anatomistas describieron tres fascículos de unión a través del túber y tallo hipofisario: el fascículo supraóptico-hipofisario, el hipotálamo-hipofisario y el túbero-hipofisario (Ingram y Ranson, 1938). En aquella época todavía existía la creencia de que la retro-hipófisis era secretora y que las fibras de aquellos fascículos transmitían el estímulo para la secreción desde el cerebro interno que constituye el hipotálamo. Tuvieron que pasar algunos años hasta que se demostrara que las supuestas hormonas del lóbulo posterior de la hipófisis (antidiurética, oxitocina y vasoprina) sólo se almacenaban en él y que su lugar de producción estaba en el propio diencéfalo, en los cuerpos neuronales de las fibras y fascículos referidos, estos últimos vehiculadores de las hormonas.

Con ello quedó firmemente establecida la noción de hormonas cerebrales y la capacidad incretora de células pertenecientes al tejido nervioso, de verdaderas neuronas, hasta entonces inimagible substrato de funciones glandulares. Pero es curioso observar la evolución que han sufrido, en los últimos años, y gracias a aportaciones decisivas, las ideas sobre la independencia neuroendocrina. A partir del descubrimiento, en 1930, por Popa y Fielding, del sistema porta en el túber y de la creencia de la función hormono-péxica a través del tallo hipofisario, se estableció el concepto de *neurocrinia* (Roussy, Mosinger, Collin); se trata de la acción de las hormonas hipofisarias sobre el diencéfalo o sistema nervioso en general. Se suponía, además, que la neurocrinia estimulaba secundariamente la increción, por parte de las células del diencéfalo *(neuricrinia)* de otras hormonas, con lo que se apuntaba a la capacidad endocrina del sistema nervioso (Scharrer).

Los estudios de estos últimos años, especialmente los acometidos por Harris, Green y otros autores, mediante distintas técnicas (trasplante de la hipófisis en ojo o riñón, sección del tallo hipofisario, destrucción selectiva de centros hipotalámicos o estímulo de los mismos, etcétera, han puesto de relieve, por una parte, el carácter inverso de la corriente sanguínea en el sistema porta, que resulta ser, por tanto, descendente y que, efectivamente, vehicula sustancias de procedencia hipotalámica y las conduce hasta la prehipófisis. Este hallazgo anatómico ha sido una demostración eficaz de un control hipotalámico humoral de buena parte de las hormonas hipofisarias del lóbulo anterior, como antes quedó establecido que las del lóbulo posterior eran, en realidad, de producción hipotalámica.

Precisamente la contribución de los doctores Guillemin y Schally se inserta en esta intrigante área de las neurohormonas diencefálicas, denominadas en un principio factores liberadores ("releasing factors"), tal vez por la repugnancia en atribuir capacidad endocrina a células nerviosas. El primer mediador bioquímico de este tipo, plenamente reconocido fue el CRF ("Corticotropin Releasing Factor"), pero en años posteriores se multiplicó con los hallazgos de ambas escuelas; se identificaron, por una parte, un factor diencéfalico para cada una de las hormonas prehipofisarias y, por otra, factores de tipo inhibidor, el primero de los cuales fue el PIF ("Prolactin Inhibitor Factor"). Los trabajos, tanto del grupo de Guillemin como del de Schally, permitieron descubrir la composición química (que se demostró polipeptídica), su peso molecular, su acción y regulación, así como otros caracteres biológicos o metabólicos, y su lugar exacto de producción en el mapa diencefálico. Su naturaleza y estructura bioquímica permitieron atribuir definitivamente el calificativo de neurohormonas a los que hasta hace poco se llamaban factores, y así también se modificó su símbolo de modo que el CRF pasó a ser CRH o ACTH-RH; la hormona liberadora de la tirotropina (TSH), antes TRF, hoy se anota en la literatura como TSH-RH y así de las demás.

El doctor Roger Louis Guillemin, de 53 años de edad, es francés de nacimiento (Dijon); estudió medicina en la Universidad de Lyon, en donde desarrolló también sus primeras funciones clínicas, docentes y de investigación. En 1949 se trasladó a Norteamérica, trabajando primero en Canadá, concretamente en el Instituto de Medicina y Cirugía experimental de Montreal durante varios años y, después, en 1953, pasó a los Estados Unidos, contratado por la Universidad de Houston (Texas), como profesor adjunto de Fisiología. Regresó a Francia, en 1960, y a los tres años, a raíz de su trabajo en el College de France, fue nombrado director de investigación en el Centre Nationale de Recherches Scientifiques de París. Al cabo de unos años volvió a los Estados Unidos, en donde actualmente reside; ha adquirido la ciudadanía estadounidense y, después de una nueva estancia en Houston, se incorporó al Instituto Salt de San Diego (California). Está casado y tiene seis hijos.

Guillemin, entre otras aportaciones, es el descubridor de la somatostatina, hormona inhibidora del crecimiento, antagonista de la STH o GH, conocida desde hace años. El y sus colaboradores lograron aislarla a partir de extractos hipotalámicos de cordero e identificaron su constitución química, como un polipéptido de 14 aminoácidos, que luego sintetizaron. Sus múltiples acciones han sido objeto de trabajo por parte de su escuela y también, independientemente, por el grupo de Schally. Guillemin ha

declarado, con motivo de la concesión del premio, que "el descubrimiento de las hormonas péptidas del cerebro permite su utilización en varios aspectos de la Medicina y algunas de estas sustancias pueden usarse para combatir la infecundidad y otras se emplean para tratar la diabetes"...; "los últimos trabajos que realicé sobre estas moléculas, en el pasado mes de mayo, se refieren a las enfermedades mentales. Es demasiado pronto para apreciar los resultados de estos descubrimientos, pero tenemos motivos para creer que serán provechosos".

El doctor Andrew Victor Schally, otro de los que comparten el premio Nobel de Medicina, ha sido un trabajador infatigable, al frente de su equipo, en la investigación endocrinológica; en particular destaca su obra pionera en el campo de las neurohormonas hipotalámicas. De su notable tenacidad y tesón da una idea el hecho de haber conseguido 500.000 hipotálamos de cerdo, recorriendo los mataderos americanos, en búsqueda de extractos en cantidad suficiente para la serie de investigaciones proyectadas: identificación del número, secuencia y reconocimiento de los aminoácidos componentes de la molécula peptídica en las distintas hormonas diencefálicas, e intentar luego su síntesis artificial, y, por otro lado, estudio de las acciones fisiológicas de cada una de ellas.

Como recuerda el doctor Gómez-Pan, médico español colaborador de Schally, éste publicó su primer trabajo en 1955, pero hasta 1969, no consiguió descubrir la estructura y la síntesis del primer péptido hipotalámico, concretamente de la hormona liberadora de tirotrofina. Sin embargo, Schally es conocido, sobre todo, por sus estudios sobre las hormonas hipotalámicas liberadoras de las gonadotrofinas. De toda su producción científica, el nuevo premio Nobel considera que lo mejor ha sido el aislamiento, reconocimiento y síntesis de este decapéptido que controla la secreción de las hormonas luteinizante y foliculoestimulante producidos en el lóbulo anterior de la hipófisis, reguladoras de toda la actividad gonadal y sexual en general. Tambien Schally trabajó en el aislamiento de la somatostatina, la neurohormona diencefálica que, como hemos dicho, inhibe la liberación de la hormona del crecimiento por parte de la prehipófisis. Es interesante señalar que los ulteriores estudios por parte de otro investigador, el inglés Hall y su grupo de Newcastle, han contribuido a demostrar los múltiples efectos de la somatostatina: inhibe también la secreción de hormona tiroestimulante (TSH), evita el desarrollo de la retinopatía y la ceguera en diabéticos, bloquea la secreción de insulina y glucagón en el páncreas y actúa incluso, directa o indirectamente, en glándulas exocrinas (en el estómago frena la secreción de ácido y de pepsina y en el intestino afecta la secreción de colecistoquinina, pancreocinina, secretina y otros mediadores bioquímicos).

Schally, que ha visitado España varias veces y que habló hace un par de años en nuestra Escuela de Endocrinología, en Barcelona, es oriundo de Wilno (Polonia), donde nació el 30 de noviembre de 1926, posteriormente adquirió ciudadanía canadiense y en la actualidad posee la estadounidense. Está casado y tiene dos hijos. Su periplo científico, en la etapa formativa y de la de investigador, transcurre por Inglaterra, Escocia y Canadá para pasar luego a Houston (Texas); hoy es jefe de los laboratorios de Endocrinología y Polipéptidos del Veterans Administrations Hospital de New Orleans, en donde también pertenece al claustro docente de la Tudane University School of Medicine.

Rosalyn Yalow, trabajando junto con Solomon Berson y otros colaboradores, han constituido la avanzadilla en la introducción de una nueva metodología analítica, que ha revolucionado en los últimos años los laboratorios hormonales, y, en general, las determinaciones de muchos compuestos biológicos que, por su pequeñísima concentración en los humores, escapaban a la detección con las técnicas convencionales. Se trata del radioinmunoensayo (RIA: "radio-inmuno-assay"), es decir, de una técnica radioinmunológica.

El fundamento de estos nuevos procedimientos analíticos parte de la obtención de un anticuerpo específico frente a la hormona que se trata de cuantificar y capaz de ligarse a ella. Luego se trata de establecer una competencia entre este anticuerpo y fracciones de la hormona; una fracción marcada con un isótopo radiactivo y otra sin marcar, que es precisamente la que se trata de determinar en un líquido biológico. Para ello, se incuban juntos los tres elementos: una cantidad conocida de antisuero específico para la hormona (que contiene el anticuerpo), otra de la hormona marcada y, la tercera, el plasma con la hormona cuya concentración en aquella hormona se desconoce. Establecido el equilibrio al cabo de unas horas, el anticuerpo se reparte entre las dos fracciones de la hormona proporcionalmente a sus cantidades respectivas. Separada la hormona ligada al anticuerpo, de la que ha quedado libre, se procede a medir la radiactividad de la primera y comparar la proporción de aquélla respecto de la segunda, lo cual depende directamente de la cantidad de hormona no marcada, la procedente del plasma problema; mediante unas curvas convencionales se puede determinar la concentración de esta última.

De esta forma se han podido determinar la concentración en sangre y orina de una serie de hormonas y otros mediadores bioquímicos: testosterona, aldosterona, insulina, ACTH, cortisol, angiotensina, antígeno Australia, etcétera. En el comunicado del Instituto Karolinska de Estocolmo, que decide el premio Nobel de Medicina, se lee: "En una serie de trabajos realizados entre 1936 y 1960, Yalow y su colaborador Solomon Berson describieron detalladamente el método de ensayos radioinmunológicos (RIA). Estos trabajos iban acompañados de una combinación espectacular de inmunología, investigaciones isotópicas, matemáticas y médicas. Asistimos al nacimiento de una nueva era en Endocrinología, iniciada por la doctora Yalow". Esta nueva Endocrinología continúa desarrollándose e ininterrumpidamente va ensanchando el horizonte de las causas y la naturaleza de las enfermedades en todos los campos médicos.

La doctora Rosalyn Yalow nació en Nueva York en julio de 1921, está casada y tiene dos hijos. Estudió en la Universidad de Illinois y pronto se dedicó allí mismo a tareas de investigación. Luego pasó a trabajar en los servicios médicos del Hospital de Veteranos de Bronx, en Nueva York, donde llegó a ser Directora del Servicio de Medicina Nuclear. En este mismo hospital sigue trabajando en la actualidad. En 1972, la Endocrine Society americana le concedió la medalla Koch junto a Berson; además de premiar sus estudios sobre el RIA, se citó su contribución al estudio de la inmunogenicidad de la insulina, la demostración de altos niveles de insulina en la diabetes de la edad madura, la aplicación del RIA al diagnóstico del insulinoma, la observación de la labilidad en la secreción de la hormona de crecimiento a estímulos metabólicos y la observación de la heterogeneidad de la paratohormona, ACTH y gastrina, entre otros descubrimientos.

La Fundación Nobel dice también de ella que "la doctora Yalow y sus colaboradores dirigieron sus investigaciones sobre la diabetes por nuevos senderos y han conseguido así una nueva dimensión. Ha sido un trabajo pionero del más alto nivel cuyos resultados han causado gran impacto". (A. B.)

Gatos y vías comerciales

Los gatos han vivido en asociación con las personas desde hace mucho tiempo, pero raras han sido las ocasiones en que éstos han sido criados por el hombre con el fin de seleccionar, según su voluntad, unas características determinadas

Neil B. Todd

A primera vista, no parece que los gatos puedan ofrecer una ayuda y un complemento significativo a la historia humana. Sin embargo, si uno se para a considerar que hace tan sólo 200 años apenas empezaban a escribirse historias aceptables de las poblaciones humanas, que la propia escritura se remonta sólo a unos 6000 años y que, en un gran número de poblaciones, los testimonios históricos, lingüísticos y culturales son incompletos o inexistentes, entonces los gatos aparecen en una nueva perspectiva. Hace mucho tiempo que viven en asociación con los seres humanos, pero nunca tuvieron ningún tipo de significado económico y sólo en raras ocasiones han importado algo desde el punto de vista social. Contrariamente a lo ocurrido con otros animales domesticados, los gatos han seguido una evolución genética esencialmente libre, dependiendo sólo de ellos mismos. Por tal razón, resulta particularmente interesante estudiar la genética de poblaciones de los gatos; no sólo nos habrá de revelar la evolución de los mismos, sino que también nos abrirá nuevos horizontes en la comprensión de los movimientos de las poblaciones humanas.

De un modo general, el estudio de la genética de poblaciones se ha convertido en un instrumento de gran utilidad para esclarecer la historia y la prehistoria humanas y, de una forma especial, para resolver problemas referentes al origen y a la dispersión de plantas y animales. No obstante, cada estudio individual pone de manifiesto las limitaciones inherentes tanto al organismo que se investiga como al enfoque que se da al estudio. Por tanto, si lo que se desea es obtener una mayor precisión en la interpretación de acontecimientos complejos, es mejor estudiar varias especies distintas. Aunque sólo fuera por este motivo, los gatos resultan los más apropiados para el estudio, pero presentan otras ventajas adicionales, además de las que ya he mencionado.

Entre las ventajas se cuenta la presencia de polimorfismos visibles, es decir, rasgos variables, muy fáciles de observar y cuyas bases genéticas se conocen relativamente bien. La mayoría de los polimorfismos están relacionados con el color, la textura o la distribución de formas y colores del pelo, características todas ellas que se pueden diagnosticar y grabar en la memoria de un simple vistazo. Por tanto, la acumulación de datos es un procedimiento relativamente sencillo y barato, que carece de embarazosas y complejas técnicas y tecnologías.

Gracias a la intensa actividad comercial y exploradora de los seres humanos, los gatos son cosmopolitas. A partir de sus diversos puntos de origen situados en el Viejo Mundo, estos animales se han propagado por todas las zonas habitadas de la tierra. Incluso persisten en lugares que el hombre abandonó. Además de hallarse por doquier, su número es muy elevado. La mayoría de sus poblaciones cuentan con un mínimo de varios centenares y no hay duda de que bastantes alcanzan varios millones. Por tanto, el genetista, cuando se dispone a estudiar los gatos, se encuentra frente a unas poblaciones de gran tamaño, con una estabilidad genética inherente.

Aunque para asentar un juicio de algunos hechos complejos serían necesarias observaciones más profundas, se puede decir que la mayoría de las poblaciones felinas se aproximan a lo que en genética se conoce con el nombre de equilibrio de Hardy-Weinberg. Esta expresión significa que se pueden establecer las frecuencias de varios factores genéticos en términos de distribuciones matemáticas teóricas, aplicables a los casos en que se supone que las poblaciones son muy grandes, que la reproducción se realiza al azar y que no se encuentran las complejidades introducidas por la mutación, la selección y la migración. Es evidente que, en el mundo real, nunca se cumplen estas condiciones. Si se cumplieran, a los genetistas no les quedaría gran cosa por estudiar. El hecho de que las poblaciones felinas concuerden aparentemente con el equilibrio de Hardy-Weinberg no significa que no sean importantes los efectos de la mutación, la selección y la migración; significa simplemente que el tiempo necesario para que estos factores provoquen cambios en la población es relativamente largo. Podemos atribuir, pues, las diferencias entre poblaciones a factores sistemáticos y no a unas fuerzas variables. En el caso del gato, una de estas fuerzas sistemáticas es el hombre; al llevarse consigo a los gatos cualquiera que fuera el lugar a donde se desplazaba, permitía la realización, sin darse cuenta, de una serie de experimentos de genética de poblaciones que ahora podemos ya evaluar.

El método básico en genética de poblaciones es el análisis de una población. En la práctica, el estudio se suele hacer a partir de un muestreo. Hay muchas formas de tomar una muestra, pero la finalidad siempre es la misma: tabular la frecuencia de los fenotipos (caracteres observables del organismo) controlados por alelos mutantes (genes variables). En las poblaciones felinas, echando un simple vistazo a los gatos, se pueden detectar simultáneamente de siete a diez caracteres de este tipo.

TRES GATOS MUTANTES componen la ilustración de la página opuesta, en la que se los compara con el tipo "salvaje", o no mutante *(a)*. Los mutantes, que presentan caracteres controlados por determinados alelos, o genes variables, son antiagutí *(b)*, cuyo tipo más frecuente es el gato negro; tabí manchado *(c)* y naranja ligado al sexo *(d)*. El color, la distribución de manchas y colores y la textura del pelaje resultan ser un medio eficaz para reconocer la presencia de los distintos alelos mutantes.


a
b
c
d

Es evidente que los gatos en sí sólo representan un empaquetamiento pasajero de información genética. Un análisis más profundo pone de manifiesto que lo que está evolucionando es el acervo de genes. Por tanto, lo que intenta hallar el genetista no son frecuencias fenotípicas sino frecuencias de alelos mutantes, que suelen diferir bastante de lo que se observa a simple vista debido a que cada rasgo se hereda e interacciona con los demás rasgos de una manera determinada. En el presente contexto estaría fuera de lugar una discusión acerca de las operaciones matemáticas por las que se transforman las frecuencias fenotípicas en frecuencias de alelos mutantes. Al final, a lo que se llega es a la construcción de un mapa de distribución geográfica en clinas, que indica la frecuencia con que aparecen determinados alelos mutantes en una región dada. A partir de entonces, el mapa se convierte en la observación más destacada. En este artículo se ofrecen tres de estos mapas; volveremos a hablar de ellos cuando tratemos de algunos fenotipos felinos.

Cualquier interpretación que hagamos a partir de los mapas de distribución geográfica en clinas y de los datos que reflejan presupone forzosamente el reconocimiento de que lo que estamos estudiando es el conjunto de las adaptaciones de un organismo que va cambiando de un nicho ecológico a otro. El gato ha pasado, o está pasando, de una vida salvaje y depredadora a un estilo de vida domesticada, en que el alimento se le suministra sin esfuerzo alguno por su parte. Está evolucionando hacia un nuevo nicho, que se enclava esencialmente en el hábitat urbano. Durante el trascurso de este proceso de domesticación, el gato está sometido a una nueva serie de relaciones inter e intraespecíficas.

¿Y qué tiene que ver este cambio de vida con los colores del pelaje? En los mamíferos, existe una relación bien definida (aunque su conocimiento sea todavía rudimentario) entre la pigmentación y las sustancias químicas endocrinas. Por ejemplo, el pigmento melanina y la hormona adrenalina tienen una vía metabólica común en la que, a partir de un precursor común, se llega a la síntesis de una u otra. Estudios relativos a varios tipos de animales han revelado que se pueden manipular algunos caracteres básicos del comportamiento, como pueden ser el miedo y la agresión, simplemente cruzando entre sí mutantes cuyo pelaje tenga colores diferentes. Hay pruebas de dichos efectos en ratas, ratones, viso-

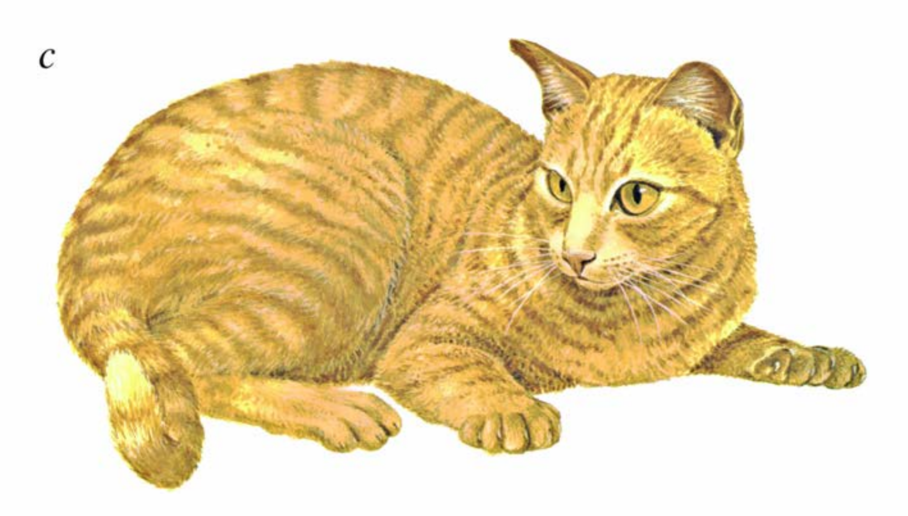
a



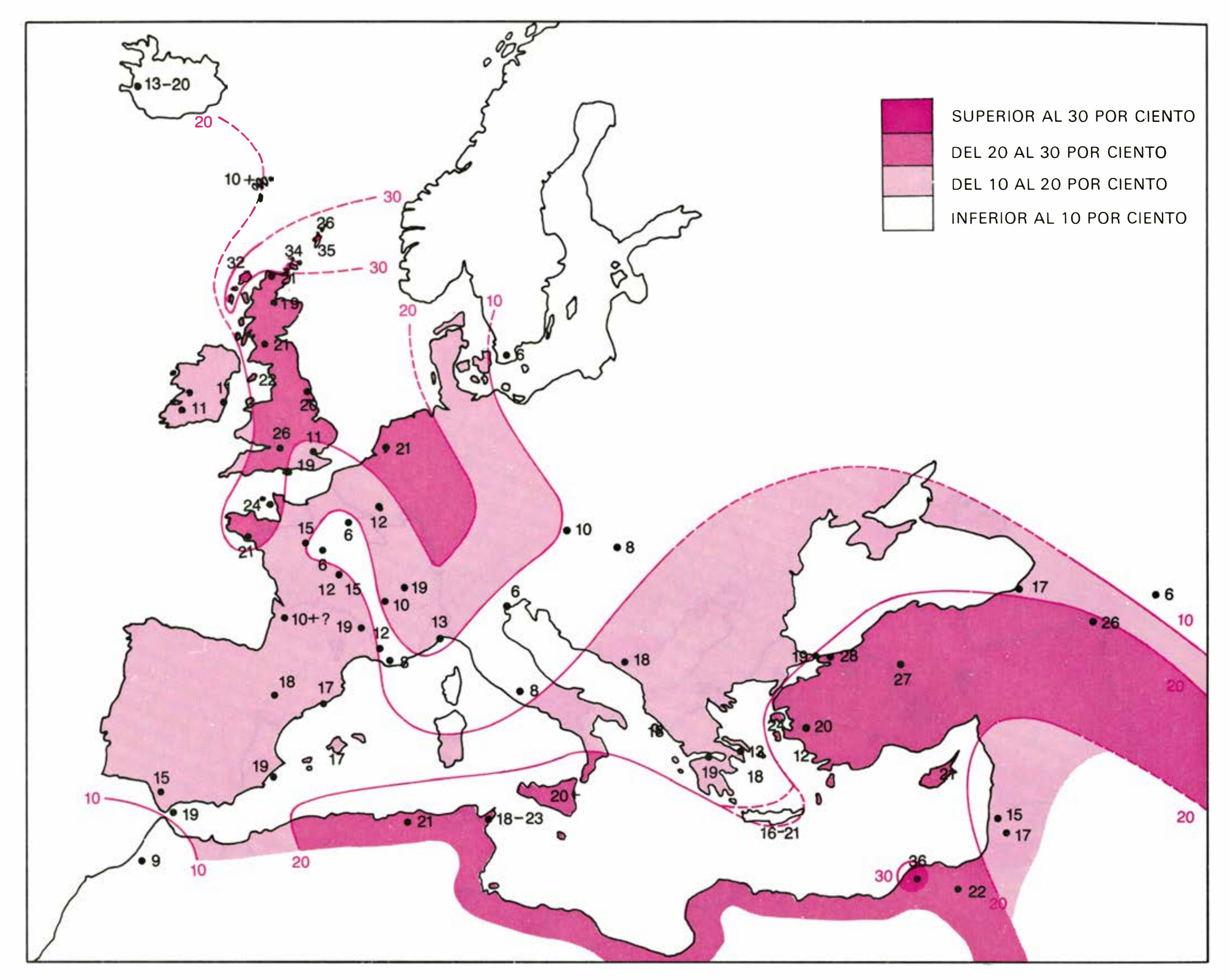
b



c



LAS VARIANTES CONTROLADAS por el alelo naranja ligado al sexo están representadas por los gatos carey *(a)*, calicó *(b)* y naranja-mermelada *(c)*. Los gatos carey y calicó suelen ser hembras porque la mutación naranja está localizada sobre el cromosoma *X* o cromosoma femenino. Una hembra tiene dos cromosomas *X*; puede expresar a la vez el carácter naranja y el carácter no naranja.



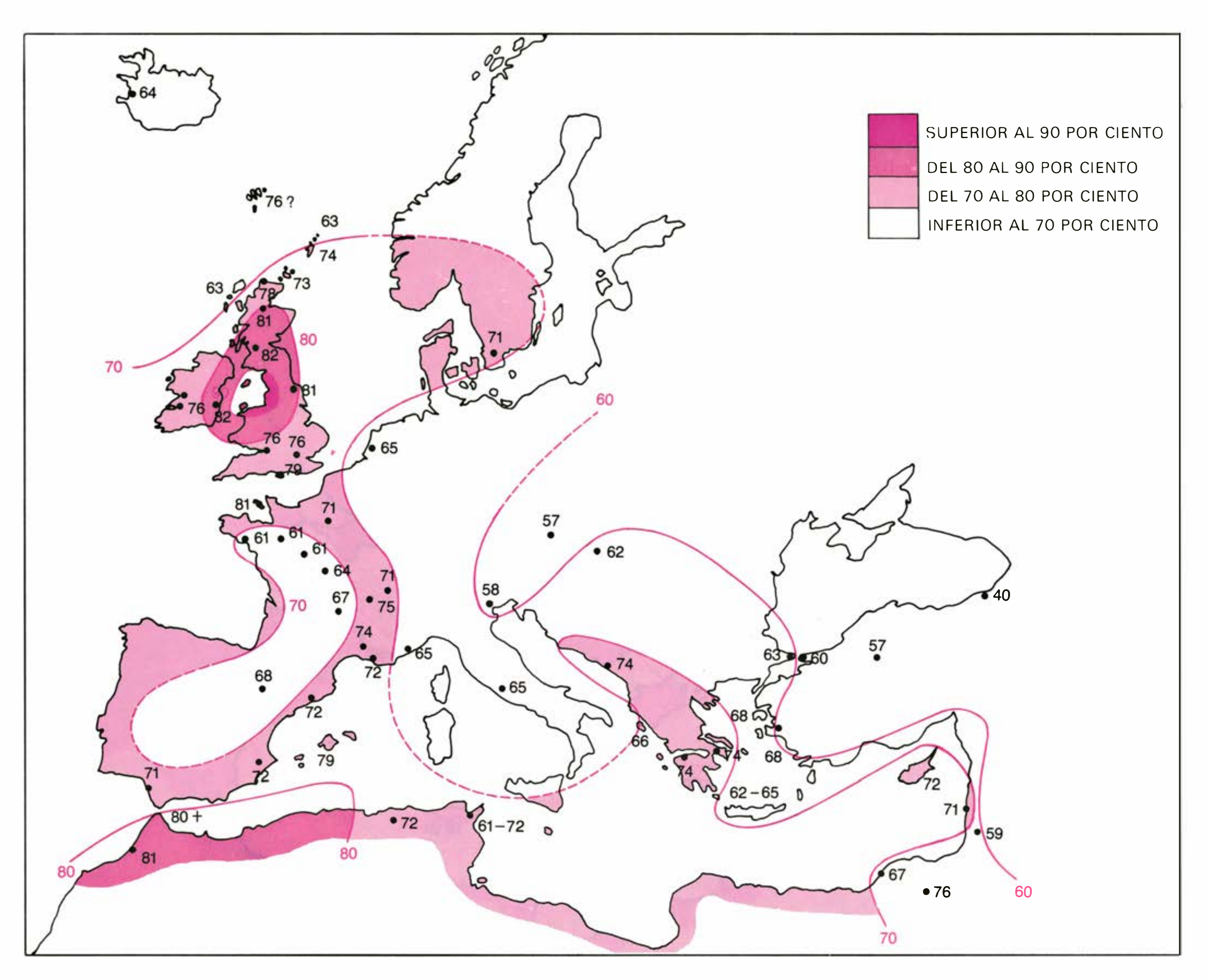
MAPA DE DISTRIBUCION EN CLINAS, que indica las frecuencias del alelo naranja ligado al sexo en términos de porcentaje. Este tipo de mapas se construyen tomando muestras de la población felina en un determinado número de localidades y traduciendo por métodos matemáticos los rasgos observables de los animales muestreados a frecuencias de alelo mutante. Las líneas continuas delimitan regiones que difieren entre sí por incrementos del 10 por ciento en la frecuencia. Se dan cifras de porcentaje más exactas para varias ciudades. Los datos se ordenan en clinas; sugieren que el foco del alelo naranja ligado al sexo (zona en que se asentó por primera vez el mutante) está en Asia Menor.

nes y zorros; si se hiciera el esfuerzo de comprobarlo, lo más seguro es que este efecto también se presentaría en los gatos y, probablemente, en todos los demás mamíferos.

Una prueba parcial, en el caso de los gatos, es que prácticamente todas las mutaciones, tomadas una a una, presentan, a lo largo de un eje urbano-rural, una selectividad diferencial en su sensibilidad a uno o varios caracteres. Por ejemplo, los perfiles genéticos de los gatos de Londres y de París presentan más semejanzas mutuas que cualquiera de ambos perfiles con el de los gatos rurales de Bretaña. Al parecer, es muy grande el poder selectivo de las fuerzas responsables de estas diferencias; en un estudio realizado en la ciudad de Glasgow, se vio que se podían distinguir perfectamente los gatos del centro de la ciudad de los que habitan en los suburbios. Este espectro de hábitats viene a ser, en muchos aspectos, una representación espacial de los cambios más importantes que han tenido lugar en las condiciones de vida de los gatos a lo largo de los siglos. Hay que recordar que tan sólo las pequeñas ciudades y comunidades rurales de hoy en día conservan las características ecológicas que fueron prácticamente universales hasta hace varios siglos. En tiempos remotos, una comunidad de 10.000 habitantes suponía un complejo urbano importante. Eran raras las ciudades, como Alejandría y Cartago, con más de 100.000 habitantes.

Puesto que las personas y los gatos están inextricablemente asociados en el proceso de domesticación, es necesario hacer un análisis de las formas de interacción que se dan entre ambas especies como también de las que *no* se dan. Ya hemos apuntado anteriormente que la dirección humana intencionada, destinada a la modificación de las frecuencias génicas de los gatos, constituye un factor despreciable en la mayoría de los casos. Sin embargo, hay algunas excepciones.

Una de ellas se podría denominar selección por novedad. Este fenómeno encuentra su origen en la tendencia humana hacia las cosas raras, curiosas o en la predilección por la extravagancia. Este rasgo quizás haya tenido implicaciones de gran alcance en la historia y el desarrollo de todos los animales domesticados, pero se hace especialmente patente hoy en día en el gato.



EL ALELO ANTIAGUTI se distribuye de la forma descrita en este mapa de clinas. El fenotipo agutí consiste en un aspecto de finas rayas claras y oscuras producido por el hecho de que cada unidad del pelaje tiene una franja de pigmentación reducida por debajo de su extremidad apical. La mutación antiagutí elimina esta franja y conduce a una pigmentación más uniforme, que suele ser negra. La distribución hace pensar en un foco principal de la mutación (hace unos 2500 años) en la parte este del Mediterráneo, que podría ser Grecia o bien Fenicia.

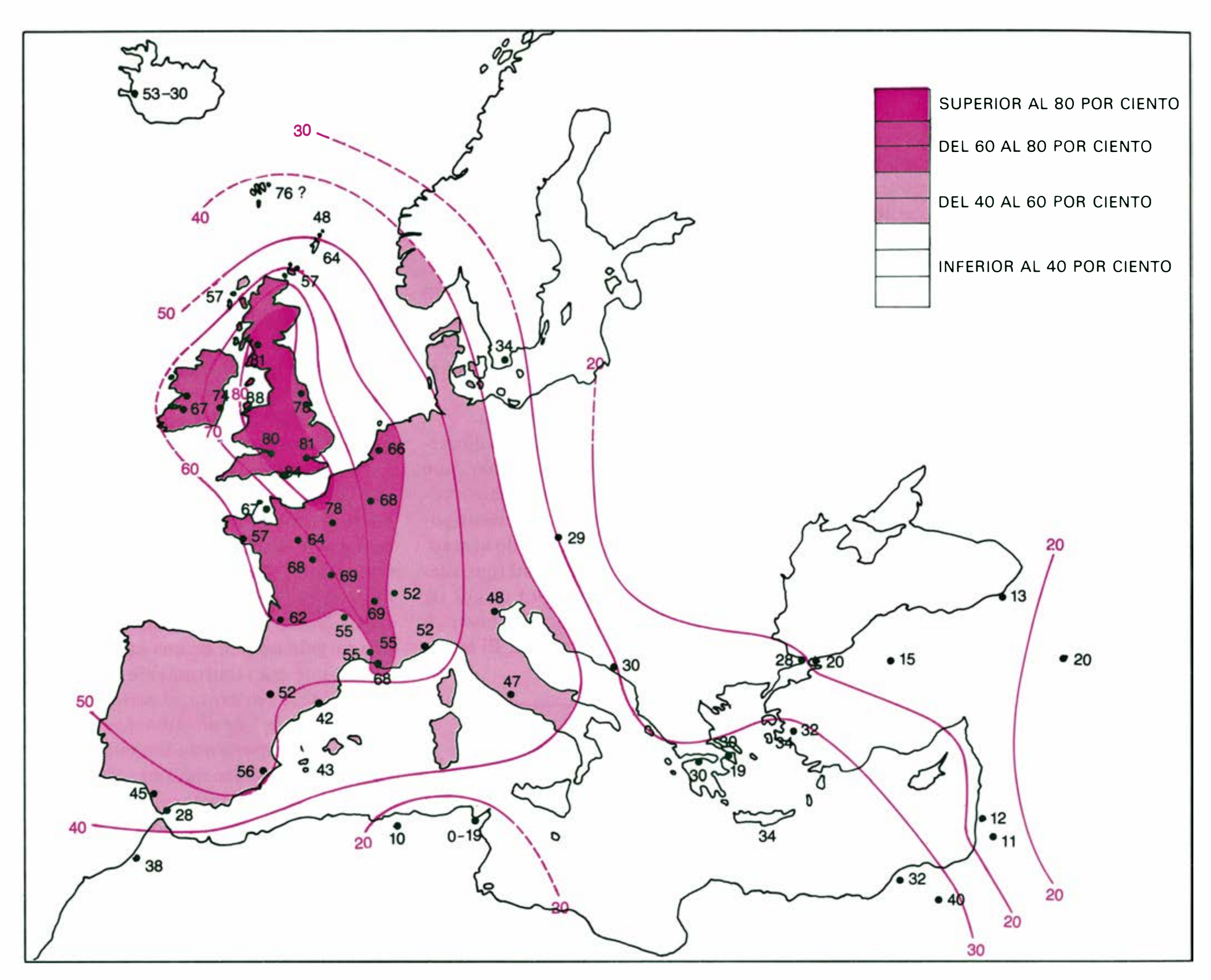
Debemos tener presente en todo momento que cada nueva mutación, cualquiera que sea su valor adaptativo intrínseco, tiene muchas probabilidades de perderse por los simples caprichos del azar.

A este proceso, los genetistas le llaman deriva genética. En las poblaciones felinas, se puede observar que varios alelos mutantes presentan valores que atribuyen a sus fenotipos respectivos una frecuencia relativa bien de aproximadamente el 10 por ciento en amplias áreas geográficas, bien un máximo de 10 por ciento en puntos dispersos. Esta cifra se refiere a los alelos mutantes heredados tanto dominantes (rasgos que controlan a los demás, como el pelaje blanco y la polidactilia, una anormalidad en los dedos de los pies, que están en número excesivo), como recesivos (pelo largo y estilo siamés) o letales (falta de cola en el gato tipo de la isla de Man), aun cuando algunos de los caracteres no resulten beneficiosos para el animal.

Parece ser que la explicación de esta curiosa distribución está en que, cuando un fenotipo alcanza una frecuencia de un 10 por ciento, ya no se considera como una novedad y, por tanto, pasa a ser objeto de cierta indiferencia. No obstante, si vuelve a disminuir la frecuencia del fenotipo, puede ocurrir que vuelva a verse favorecida.

Por consiguiente, nos hallamos en presencia de un tipo de selección digno de observar, que se podría denominar selección dependiente de la frecuencia; éste opera de suerte que facilita la introducción y el mantenimiento de la variabilidad genética en las poblaciones. Es muy probable que, en contrapartida, esta misma fuerza impida que una mutación, que resultaría altamente ventajosa, desplace totalmente al alelo opuesto (es decir, que llegue a fijarse). Este sistema de selección poco habitual quizá sea la causa última de la riqueza de variabilidad genética observada en los animales domésticos, aunque, en las especies más explotadas, pronto se dividen las variantes en razas que se van adecuando por selección a alguna necesidad humana. En el caso de los gatos, dado que estos animales no ocupan lugares de alto nivel en los asuntos económicos o sociales, estas variantes simplemente se van dispersando a través de las poblaciones.

La segunda excepción es el resultado de lo que podríamos llamar migración selectiva. Paradójicamente, lo que para la mayoría de los animales son barreras



EL ALELO TABI MANCHADO presenta esta distribución. Se puede interpretar el mapa suponiendo que hay dos focos del alelo, uno (que ocupa el lugar más avanzado) en Gran Bretaña y el otro en el nordeste de Irán. El flujo que procede del foco iraní y se dirige hacia el oeste entra en contacto con el que viene del foco europeo y se mueve hacia el este a nivel de Turquía. La hipótesis de que Gran Bretaña sea un foco del alelo se confirma por las frecuencias observadas (de portadores del alelo) en las colonias de Nueva Inglaterra, Canadá, Australia y Nueva Zelanda.

acuáticas se convierten, para los gatos, en verdaderas vías de desplazamiento.

Es evidente que los animales domésticos, en su mayoría, han sido transportados por vías acuáticas, pero este desplazamiento suele estar sometido a un control riguroso. En cambio, los gatos establecen su resistencia en las embarcaciones, van y vienen a voluntad (con el riesgo de quedarse en tierra al zarpar). Parece ser que, tan pronto como el hombre dominó el arte de la navegación comercial, los gatos se convirtieron en miembros habituales de la tripulación.

Es posible que la elección de compañeros de crucero felinos no tenga otro efecto más que el de extender la selección por novedad. Por otra parte, quizá sea promotora de un cambio genético ulterior debido a circunstancias que permitan el establecimiento de preferencias básicas estéticas (o de otro tipo) como agentes selectivos de gran eficacia. Cualesquiera que sean los motivos seguidos para la elección de los gatos, su efecto en las migraciones marítimas puede ser particularmente marcado. A menudo ocurre que los animales implicados proceden de las ciudades y, al ser transportados a una comunidad menor, pierden la inercia que representaba ser parte integrante de una gran población. Así, una preferencia humana que hubiera quedado enmascarada en una población felina estabilizada tiene la ocasión de hacerse patente en una población migratoria.

Es inevitable que algunas mutaciones confieran ventajas especiales en determinadas circunstancias. Un ejemplo familiar es el del alelo para el melanismo en las mariposas nocturnas; en circunstancias normales, es deletéreo porque hace que los insectos sean más fáciles de ver para los pájaros, pero, en las áreas industriales ennegrecidas, resulta ventajoso el color oscuro porque queda más disimulado.

Cuando la frecuencia de mutación de fondo se acopla a la selección por novedad y a la selección por migración, el resultado es una aceleración de la generación de variabilidad y una facilitación del proceso de adaptación. El mutante, que se vio previamente conservado y socorrido por una selección especial, ahora está sometido a la selección natural, en el sentido de que su destino ya no se halla directamente ligado a la idiosincrasia del hombre. La mutación puede aportar alguna superioridad intrínseca, tal co-

mo una mayor resistencia a las enfermedades o una tolerancia más amplia en cuanto a necesidades nutritivas, o puede modular el comportamiento de forma tal que queden reducidas las fricciones entre especies o entre individuos de una misma especie. Sea cual sea la contribución de la mutación, el caso es que entran en juego nuevas fuerzas que dependen de la naturaleza de la diferencia genética.

En algunos casos, los alelos mutantes resultan más beneficiosos en el estado heterocigoto. (Todo organismo recibe la mitad de sus genes de su madre y la otra mitad de su padre. Si, para un carácter dado, como puede ser el color del pelo, los dos gametos, o células sexuales, que se fusionaron durante la fecundación llevan el mismo gen, el descendiente resultante es homocigoto. Si los gametos llevan genes distintos correspondientes al mismo carácter, el descendiente resultante es heterocigoto.) La heterocigosis puede dar lugar a una "superioridad de los híbridos", en cuyo caso los mutantes alcanzarán un equilibrio que dependerá de la eficacia relativa de las dos posibilidades homocigóticas. En otros casos, los factores resultan más ventajosos sólo en una de ambas combinaciones homocigóticas. Entonces, los mutantes van siendo cada vez más abundantes hasta que llegan a fijarse, proceso cuya velocidad depende de los valores adaptativos relativos de los diversos genotipos. Finalmente, puede ocurrir que una mutación no confiera ninguna ventaja intrínseca, pero que, debido a su novedad, sea conservada y dispersada con frecuencias bajas.

De los cuatro genotipos felinos que seguidamente comentaremos, hay uno que, por lo que nos indican los hechos, confiere una superioridad a los heterocigotos: es el genotipo naranja ligado al sexo. Hay otro genotipo, el antiagutí (que suele ser negro), que manifiesta rasgos de superioridad heterocigótica u homocigótica según las circunstancias. El tercero, tabí manchado, está tan favorecido en el estado homocigótico que está desplazando de forma inexorable al alelo opuesto, es decir, al tipo salvaje. El cuarto genotipo, la polidactilia, persiste sólo como una curiosidad. Las hipótesis que hemos emitido respecto a las fuerzas selectivas se basan en parte en las distribuciones geográficas que presentan unas clinas patentes y, en parte, en otro tipo de pruebas y consideraciones teóricas. De todos modos, aún no se está en poder de pruebas inequívocas; por esta razón, la discusión que a continuación presentamos sirve principalmente para describir algunas de las hipótesis que se están proponiendo como línea de trabajo para estudios ulteriores.

Vale la pena apuntar de una forma explícita la suposición que se halla subyacente en el estudio de estos cuatro genotipos felinos: la de que las distribuciones heterogéneas actuales de los cuatro mutantes (como también de otros mutantes) han surgido a partir de una población ancestral que era relativamente uniforme y homogénea. Por tanto, el período y el lugar de "origen" de un mutante son factores de una importancia fundamental, en cuanto a su influencia sobre la distribución del mutante, factores con los que se quiere significar el tiempo y la localidad en que el mutante se asentó con firmeza por primera vez. Dicho punto concreto se transforma posteriormente en un foco más amplio a partir del cual hay dispersión hacia otros lugares y esto ocurre en todos los casos, sea cual sea el punto geográfico preciso en el que se originó el mutante. Nosotros utilizaremos el término foco para referirnos a la situación global, quedando entendido que el sitio de origen exacto puede coincidir o no con el foco.

Naranja ligado al sexo es el gen mutante responsable de una gran variedad de fenotipos conocidos, que dependen del sexo y del genotipo del individuo portador del gen y también de ciertas interacciones con otros mutantes. Los ejemplos más familiares son los gatos del tipo calicó, carey y naranja-mermelada. Los gatos carey y calicó casi siempre son hembras, ya que el fenotipo depende de la heterocigosis. (La mutación naranja se localiza en el cromosoma femenino, o cromosoma *X*. Y puesto que una hembra posee dos cromosomas de este tipo, puede ser portadora de naranja y de antinaranja y expresar simultáneamente ambos rasgos. Normalmente, los machos sólo tiene un cromosoma *X*, siendo *Y* el otro cromosoma sexual, y, en consecuencia, un macho puede expresar naranja o anti-

EL TIEMPO Y LA FRECUENCIA del alelo tabí manchado se representan aquí en un sistema de coordenadas. La fecha asignada a cada comunidad *(puntos negros)* y a cada región *(puntos coloreados)* es el año aproximado en que el lugar fue colonizado por los británicos. La frecuencia actual del alelo en cada lugar constituye una aproximación de su frecuencia en la población felina británica en el tiempo de colonización. La línea que representa la media predice con exactitud la frecuencia actual del alelo en Gran Bretaña *(parte superior derecha)* lo cual confirma la suposición hecha a partir del mapa de clinas de tabí manchado; en efecto, el alelo ha aumentado de frecuencia en poco tiempo en Gran Bretaña y se va difundiendo con fuerza por todas las áreas circundantes.

naranja, pero no los dos a la vez.) Estos detalles constituyen por sí solos un tema apasionante, pero lo que resulta interesante en nuestro contexto es que la frecuencia del alelo naranja ligado al sexo no supera el 36 por ciento en ninguna parte del mundo; el nivel habitual es próximo al 25 por ciento. Esta observación nos conduce a una conclusión provisional: la de que el equilibrio básico se encuentra alrededor de 25 por ciento, lo cual es un nivel que concuerda bastante con lo que se espera obtener en los casos de superioridad de los heterocigotos. Por tanto, la interpretación menos osada de la distribución en clinas del alelo naranja ligado al sexo es que, en las regiones donde el porcentaje es inferior al 25 por ciento, aún no se ha alcanzado un equilibrio y, en donde es superior, se ha alterado el equilibrio.

El mapa de clinas que muestra la distribución del alelo naranja ligado al sexo presenta una forma curiosa: un área central irregular con frecuencias notablemente menores que las que se hallan fuera del centro. No se puede dar más que una explicación vaga a esta distribución; si bien, al parecer, puede explicarse una de las características geográficas: el pasillo que atraviesa Francia, uniendo Londres con el Mediterráneo. Los valles del Sena y del Ródano, ampliados en los últimos años por un sistema elaborado de canales para barcazas, han constituido, desde hace mucho tiempo, una ruta marítima muy desarrollada que puede servir como alternativa al transporte de mercancías por la ruta del estrecho de Gibraltar.

Esta vía tierra adentro es una característica constante en la distribución de tres de los mutantes que estamos considerando. Fue un modo de dispersión clarísimo para el antiagutí, el tabí manchado y para otros mutantes, pero la influencia que ejerció sobre naranja ligado al sexo fue, al parecer, de signo completamente opuesto. Los valores bajos de mutantes en las partes centrales, en particular Roma y Marsella, y quizá también Londres, se podrían atribuir a la presencia larga y continuada de grandes poblaciones de gatos. El alelo naranja ligado al sexo, al alcanzar estas áreas, sólo representaría en ellas una pequeña fracción de todos los alelos presentes. Por tanto, la baja frecuencia del mutante en el pasillo refleja únicamente que las posibilidades de que los gatos naranja ligado al sexo se recluten como emigrantes son muy reducidas. Por otra parte, es probable que la larga estabilidad de la población felina sea un reflejo de los siglos de hegemonía romana en esta región.

¿Se podría deducir el foco de origen del mutante a partir de la distribución en clinas? La frecuencia relativamente alta de naranja ligado al sexo a lo largo de la costa norte de Africa, en las islas Baleares y a lo largo de la costa mediterránea española hace pensar en una migración procedente del este y llegada por vía marítima. Así se nos presenta Asia Menor como un foco potencial, ya que en aquella región las frecuencias son uniformes, en comparación, y es muy probable que hayan alcanzado un equilibrio en un área muy extensa. Con alguna excepción, los niveles disminuyen en todas las direcciones.

Pero entonces, ¿qué es lo que explica la elevada frecuencia del mutante en las islas situadas frente a las costas del norte y del oeste de Escocia y, en menor grado, en las zonas rurales de Islandia, en las islas Føroyar y en la de Man? La frecuencia de alelo blanco dominante también es, en estas áreas, superior a la de cualquier otro lugar. Pudiera deberse ello a que, unos 1000 años después de haberse introducido en estos puestos más avanzados de la civilización europea, los gatos fueran aún el reflejo de lo que podían constituir las preferencias estéticas de los vikingos.

La única región, aparte de las mencionadas, donde se da esta combinación de frecuencias de naranja y blanco dominantes es, según se ha comprobado, en el este de Turquía, en el lejano departamento de Van. Vale la pena resaltar que los gatos blanco dominante, aunque pueden resultar agradables a la vista, constituyen un genotipo que está básicamente en desventaja respecto a otros. Estos gatos tienen una viabilidad reducida y los que sobreviven suelen padecer defectos de la vista y del oído y son propensos a determinados tipos de cáncer de piel. En cualquier sitio donde se hallen frecuencias altas de este genotipo, hay que reconocer que se debe a la mano del hombre. Los hechos nos sugieren que los vikingos transportaron de forma selectiva este perfil naranja ligado al sexo y blanco dominante desde sus puntos de contacto con el Mar Negro hasta las zonas más occidentales, implantando estos mutantes en el norte del Atlántico.

En cuanto al mutante antiagutí, digamos que la manifestación más común de este alelo es el pelaje negro. Citemos algunos rasgos generales. El carácter agutí (agutí es el nombre de un roedor sudamericano que constituye un ejemplo del conjunto de las características "agu-

tí") viene a ser un rasgo básico de los mamíferos. El fenotipo agutí se debe primariamente a la distribución de la melanina en el pelo, de la que resulta la producción de una franja de reducida pigmentación por debajo de la extremidad de cada unidad pilosa. El aspecto externo resultante es un pelaje de finas rayas claras y oscuras, como una mezcla de sal y pimienta. Una mutación recesiva sencilla, antiagutí, elimina esta franja y, por tanto, da lugar a un pelaje más o menos uniforme en cuanto a la pigmentación a lo largo de cada unidad de pelo.

La mayoría de los mamíferos son agutí; en cambio, la mayoría de los mamíferos domesticados son antiagutí. En este aspecto, el gato constituye una excepción, pues en cualquier lugar presenta un dimorfismo acentuado con la presencia de los dos estados opuestos. El estudio de 150 poblaciones pone de manifiesto que antiagutí se halla por encima de un 50 por ciento. Este hecho es una fuerte prueba circunstancial de que existe un equilibrio básico en este nivel y de que probablemente venga determinado por una ventaja de los heterocigotos.

¿Hay algún reflejo de la mano del hombre en la distribución de los gatos antiagutí? Las frecuencias más altas del mutante se encuentran en Gran Bretaña y en algunas partes del noroeste de Africa. Se observan frecuencias ligeramente menores con predominio en el estrecho pasillo central de Francia y por toda la zona mediterránea; Roma y Venecia constituyen excepciones muy evidentes.

Parece bastante claro que el mutante antiagutí está favorecido por el hábitat urbano. Si se interpretan los hechos como significativos de que el mutante es un instrumento que fomenta la asociación gatos-humanos, se puede llegar a conclusiones muy interesantes. Los valores altos de Gran Bretaña encuentran su origen en el alto grado de asociación implícito en la migración selectiva, reforzado por una urbanización intensiva. El significado del pasillo que atraviesa Francia, en este caso, también es que tenía un papel importante en el movimiento del comercio.

Aunque no hay duda de que Inglaterra y el noroeste de Africa están hoy actuando como focos principales, las elevadas frecuencias del mutante que allí se encuentran se crearon inicialmente por una intensificación secuencial de las frecuencias mediante la migración selectiva. En otras palabras, cada vez que la gente se embarcaba para expander su esfera de actividad civilizadora, se llevaba consigo una muestra de concentración creciente de gatos antiagutí. ¿Cuál fue el primer paso de este proceso en la región mediterránea clásica? ¿Lo dieron los griegos y los fenicios al colonizar Massilia y Cartago? ¿Fueron los cartaginenses los que dieron el segundo paso al fundar comunidades en las islas Baleares, en Algeciras, Tánger y Rabat?

El alelo tabí manchado constituye una aproximación muy cercana a lo que sería la diseminación ideal de un mutante. La simple observación del mapa de distribución revela claramente un foco británico (y en este caso probablemente un origen británico primario), una penetración en el Mediterráneo a través de Francia y una expansión marítima rápida en dirección al este. Se ha identificado recientemente un foco en el este, en el interior del nordeste de Irán o en zonas adyacentes. El flujo que parte de este foco en dirección al oeste parece ponerse en contacto con el flujo procedente de Europa en la meseta de Anatolia, de Asia Menor.

La ventaja selectiva que presenta el mutante tabí manchado sigue siendo un misterio. Está dispersándose claramente como una epidemia que, por lo que parece, acabará englobando todas las poblaciones de gatos. Incluso entre los gatos salvajes de Tasmania, el mutante resulta ser superior al alelo opuesto (el tipo salvaje), lo que demuestra la existencia de una presión de selección que está reduciendo la mayoría de los demás mutantes a niveles muy bajos.

Ya que parece probable que el alelo tabí manchado apareciera recientemente, y teniendo en cuenta la fuerza de dispersión que lo impulsa, son pocas las pistas que podemos obtener para trazar su historia en las clinas del Viejo Continente. Sin embargo, los datos del Nuevo Mundo, de Australia y de otros lugares pueden servirnos para completar la imagen de Gran Bretaña como foco.

A partir del siglo XVII, se empezaron a exportar los gatos británicos a tierras en las que no había gatos indígenas. Tanto Nueva Zelanda como Canadá, Australia y Nueva Inglaterra fueron zonas receptoras, en distintas épocas, de los propágulos británicos. Si se acepta la premisa de que cada una de estas poblaciones se fundó a partir de una muestra representativa de la población de los gatos británicos, tomándose cada muestra en un tiempo distinto a lo largo de un período de 300 años, se establece una relación interesante entre la frecuencia de determinados alelos y el momento concreto del muestreo. El tipo tabí manchado es el

alelo clave para este análisis, ya que pone de manifiesto una correlación muy alta.

Supongamos que a los distintos lugares receptores se les asignan unas fechas que concuerdan con el momento de la colonización inicial y que se toman los perfiles genéticos actuales de los gatos de esas áreas para deducir los de la población felina original; se puede representar luego el tiempo en relación con la frecuencia. Esta gráfica revela una tendencia, establecida en los primeros 200 años, que permite predecir la frecuencia actual del tabí manchado en Gran Bretaña con una precisión muy alta. Este hecho da mayor fuerza a nuestra consideración, basada en el mapa de distribución del Viejo Continente, de que el alelo tabí manchado ha ido aumentando de forma rápida en Gran Bretaña y dispersándose hacia las poblaciones circundantes. Probablemente, análisis similares realizados en relación con la exploración y colonización de los holandeses, franceses, portugueses y españoles darían resultados similares.

Mi último ejemplo se refiere a una mutación que se manifiesta por un número suplementario de dedos. (Polidactilia significa muchos dedos; no es, efectivamente, un término adecuado.) Los gatos que se caracterizan por esta condición abundan en Nueva Inglaterra, Nuevo Brunswick y Nueva Escocia, zonas en que pueden representar más del 10 por ciento de todos los individuos. En cambio, son muy poco frecuentes en otras zonas.

Los hechos nos sugieren que el lugar de origen del mutante fue Nueva Inglaterra, o por lo menos un foco de Nueva Inglaterra, a partir del cual se dio la dispersión de los gatos, probablemente por la novedad que suponían. Basándonos en esta suposición, más lógico sería detectar la presencia de estos gatos, por ejemplo, en Boston hacia mediados del siglo XVIII. La hipotesis queda confirmada por el hallazgo en Halifax de una frecuencia de estos gatos igual a la de Nueva Inglaterra. La ciudad de Halifax no se fundó hasta mediados del siglo XVIII; posteriormente se destacó como centro de actividad comercial relacionado con Nueva Inglaterra y, más tarde, constituyó sede del refugio para miles de súbditos fieles a Gran Bretaña en el momento de la rebelión americana. Al parecer, la gente que huyó y se refugió en Halifax (Nueva Escocia) se llevó consigo algo más que sus pertenencias, su profesión y sus convicciones políticas. También se llevaron sus gatos polidáctilos.

Los usos de los instrumentos de sílex del Paleolítico

El examen microscópico de los filos trabajados de ciertos utensilios líticos, empleados por los antiguos cazadores, permite diferenciar los que se usaron para raer pieles, cortar carne y serrar madera

Lawrence H. Keeley

Casi la única prueba de la presencia del hombre sobre la tierra durante un período de más de medio millón de años está constituida por innumerables instrumentos de piedra. Algunos están fabricados en basalto, otros en cuarzo o cuarcita y otros en obsidiana. En muchas localidades, la mayoría están hechos de sílex. Tan pronto como se reconoció que estos objetos habían sido fabricados por el hombre se les dieron nombres basados en conjeturas sobre sus probables usos. Comenzaron los franceses con la palabra *coup-de-poing,* cuya versión española es bifaz o "hacha de mano" (*hand axe,* en inglés), siguiendo a éste otros muchos nombres funcionales: "raederas", "hoja", "punta", "buril" o "cincel", etc. Aunque estos términos han sido utilizados por generaciones de prehistoriadores, apenas si se ha obtenido alguna prueba tangible sobre el uso a que realmente estaban destinados estos instrumentos de piedra.

En los últimos quince años, los paleoantropólogos han decidido cambiar el estado de cosas investigando sobre estas cuestiones. El resultado ha sido el desarrollo de una metodología conocida como análisis de microdesgaste, con la que se pone de manifiesto el uso dado a muchos utensilios primitivos de sílex. El testimonio queda casi indeleblemente registrado en forma de microscópicas marcas de desgaste sobre el filo trabajado del sílex.

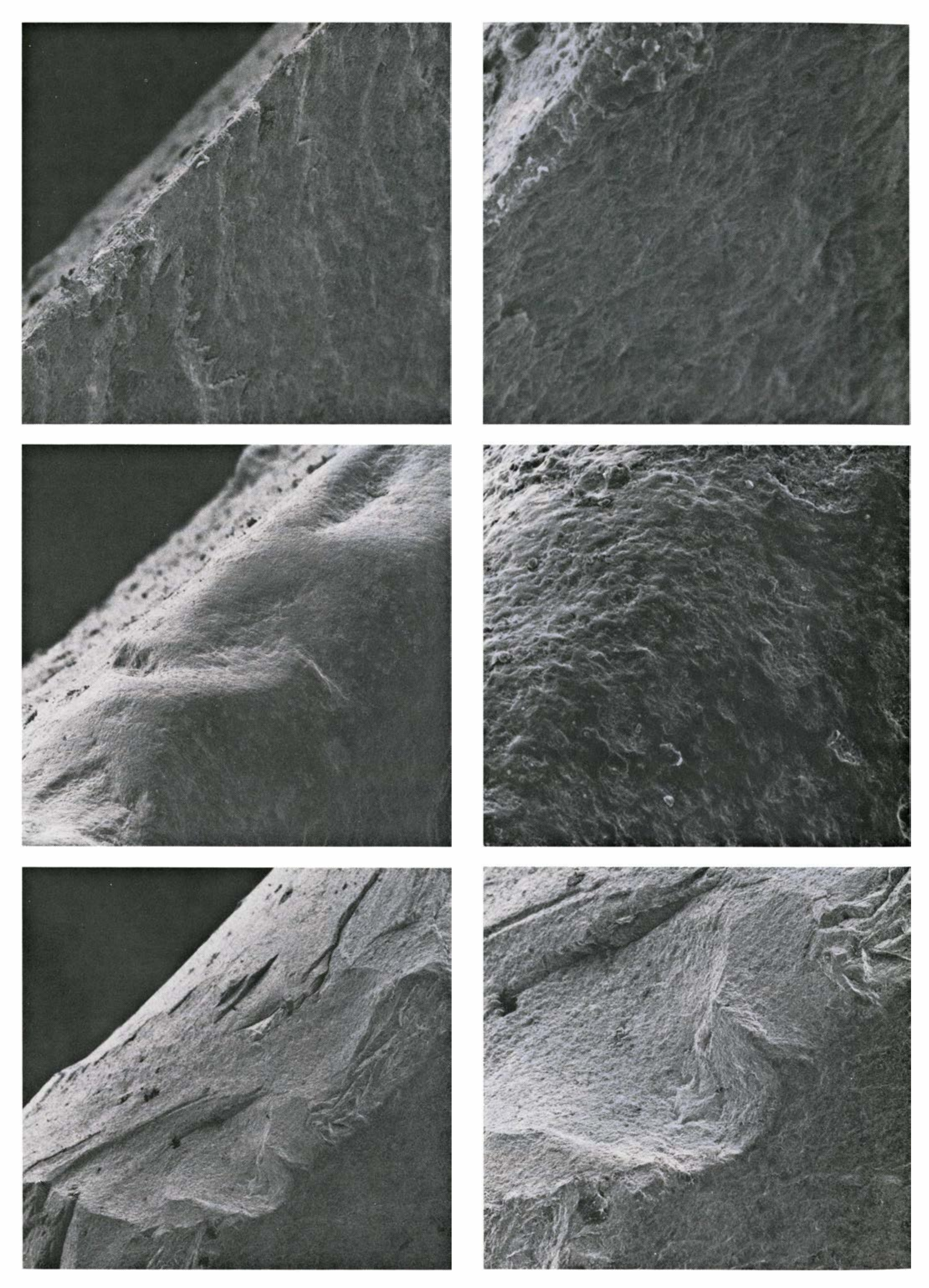
Una razón del interés general despertado repentinamente sobre el uso de los instrumentos líticos se debe a que el progreso en los métodos de datación absoluta, como el análisis del carbono-14, ha resuelto dos antiguas preocupaciones de muchos prehistoriadores. La primera era el establecimiento de cronologías relativas, en ausencia de datación absoluta. La segunda estaba íntimamente relacionada con la primera: investigar las semejanzas "culturales" entre asociaciones de utensilios de piedra procedentes de diferentes zonas. Tales semejanzas ayudan a la hora de construir cronologías regionales interrelacionadas. Al comienzo de los años sesenta, una nueva escuela de prehistoriadores comenzó a proponer nuevas hipótesis para explicar las variaciones que se registraban entre los conjuntos regionales de utensilios así como en el seno de los mismos.

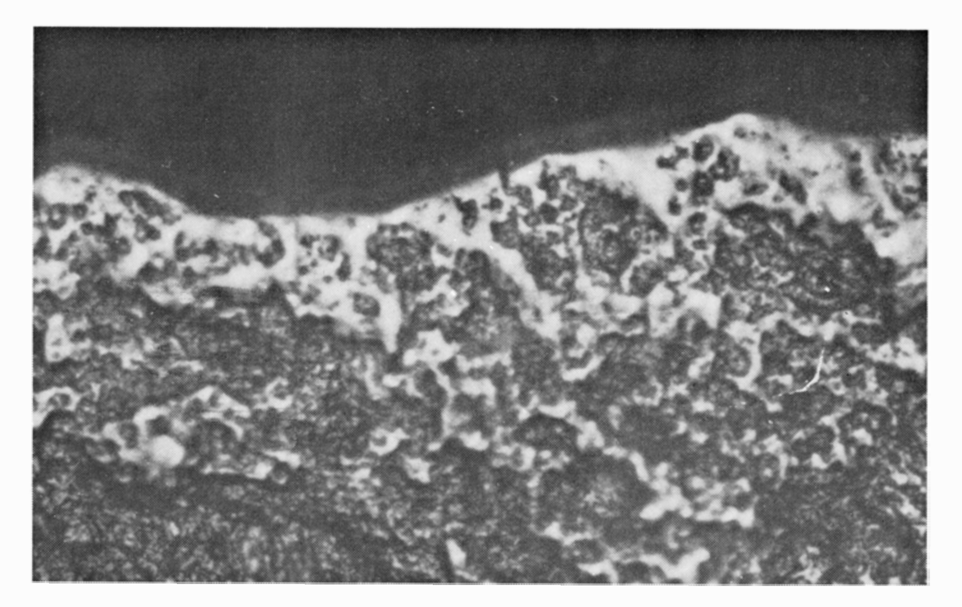
Bajo su punto de vista, las variaciones se atribuyeron, más que a diferencias cronológicas y culturales, a diferencias funcionales o de uso. Así, la nueva escuela trata de explicar las diferencias entre las clases de utensilios presentes en dos asociaciones aproximadamente contemporáneas en términos de tipos de actividad diferentes que los usuarios de los instrumentos podrían haber ejercido en ambos lugares. Los prehistoriadores de esta escuela objetaron que, al atribuir tales diferencias a distinciones "culturales" entre dos comunidades aisladas, la vieja escuela estaba interpretando defectuosamente los testimonios.

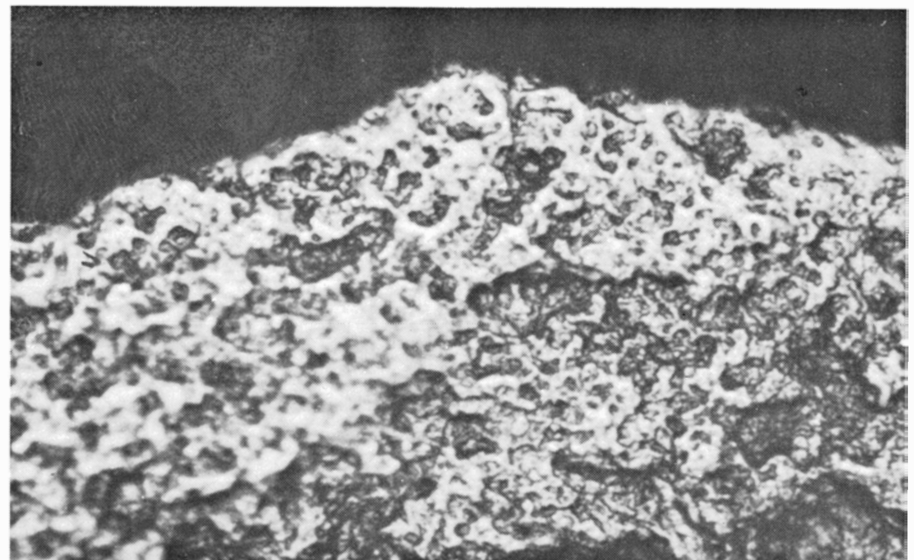
Un requisito previo fundamental para someter a prueba la hipótesis funcional o de uso era un conocimiento detallado de los artefactos utilizados y cómo y para qué se habían usado. En 1964, se publicó *Phehistoric Technology,* versión inglesa de un compendio de estudios sobre el uso de instrumentos, dirigido por el prehistoriador ruso S. A. Semenov. Este científico y sus colegas de la Academia de Ciencias de Leningrado habían comprobado que los utensilios de piedra, por dura que ésta fuese, conservaban verdaderas marcas reales de su uso en forma de pulimentos, estriaciones y otras alteraciones sobre el borde trabajado de los instrumentos. En la mayoría de los casos, las marcas de desgaste eran visibles únicamente a aumentos muy altos. Los partidarios de la nueva escuela, de Inglaterra y Estados Unidos, barruntaron que, por fin, tenían a su alcance la forma de buscar el tipo de información adecuada sobre el uso de instrumentos líticos que necesitaba la nueva hipótesis.

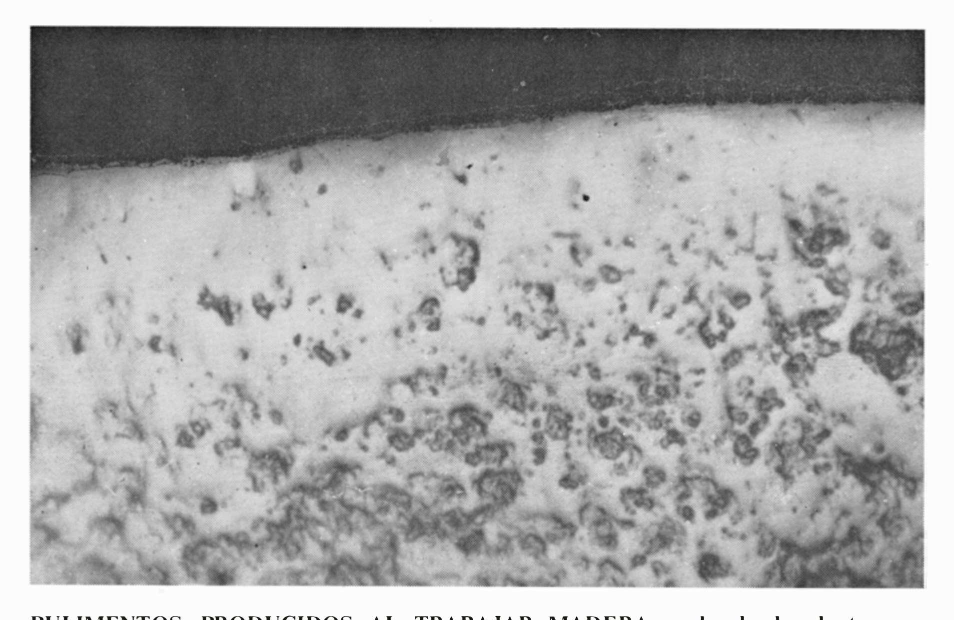
Las interpretaciones funcionales de Semenov sobre los usos de los utensilios de piedra paleolíticos y posteriores, descubiertos en la Unión Soviética, eran fascinantes pero no totalmente asequibles, ya que no había incluido en su publicación un informe detallado de la metodología que constituía el fundamento de sus interpretaciones. Para

MICROTOPOGRAFIA ALTERADA de las superficies trabajadas de instrumentos de sílex *(página opuesta)*, **según puede observarse en las microfotografías electrónicas de barrido. En la parte superior izquierda, se ve, aumentado 300 veces, el borde de un sílex no utilizado. A su derecha, una vista más próxima de una superficie de sílex no utilizada, aumentada 1700 veces. En el centro, el borde del sílex y su superficie (aumentados 140 y 1700 veces, respectivamente) una vez empleados por el autor para raspar piel seca. El borde cortante está marcadamente redondeado, y la topografía de la superficie utilizada se ha alterado por el contacto con la piel hasta adquirir una textura extremadamente mate. En la parte inferior, se registra el borde y la superficie (aumentadas 70 y 130 veces, respectivamente) utilizados por el autor para raspar hueso. Tanto el borde como la superficie muestran la topografía irregular producida por este tipo de trabajo; sobre la superficie del sílex se han desarrollado unos hoyuelos característicos.**









PULIMENTOS PRODUCIDOS AL TRABAJAR MADERA en los bordes de tres copias modernas de instrumentos de sílex. El autor raspó madera de tejo con el instrumento que aparece en la parte superior; sobre las partes en relieve del sílex ha comenzado a aparecer un pulimento originado por madera, característicamente brillante. Con el instrumento de la parte central cortó madera de abedul; la extensión del pulimento puede atribuirse, en parte, al contacto más amplio entre el artefacto y la pieza trabajada. Con el útil de la microfotografía inferior raspó madera de tejo por un tiempo más prolongado; las depresiones del filo están borradas.

empeorar las cosas, las características particulares del equipo microscópico empleado por Semenov sólo estaban disponibles en aquel entonces en la Unión Soviética y, además, el traductor había omitido la mayor parte de los escasos detalles técnicos que Semenov había incorporado en su original.

Como consecuencia, bastantes prehistoriadores no rusos comenzaron a trabajar sobre análisis de microdesgaste provistos únicamente de la traducción del libro de Semenov y microscopios estereoscópicos que muchas veces sólo tenían un aumento máximo de 80 diámetros. A estas desventajas técnicas había que añadir que los utensilios investigados por estos científicos estaban hechos de un material lítico distinto del de los instrumentos encontrados en la Unión Soviética. La desilusión y el desengaño embargaron a un investigador tras otro cuando hallaron que los resultados de Semenov no podían corroborarse.

Este estado de cosas era, sin embargo, bastante previsible. En primer lugar, en la mayoría de los casos los investigadores ni siquiera habían podido ver huellas de microdesgaste, tales como los pulimentos y estriaciones que había observado Semenov, debido a que los aumentos con que habían trabajado eran demasiado bajos. En segundo lugar, los rasgos de desgaste poco aumentados que podían observar (deterioro primario del filo; pequeñas señales de fisuras y descascarillamiento sobre el filo trabajado del utensilio) no permitían interpretaciones precisas e inequívocas del uso de los instrumentos. Muchos investigadores llegaron a la conclusión de que las interpretaciones de Semenov eran meras conjeturas, y que el simple análisis de microdesgaste no resolvía el problema. No obstante, la solicitud de información sobre los usos de los instrumentos de piedra demostraba que el interés sobre la investigación no había decaído.

Como resultado de esta serie de hechos, los científicos no rusos concentraron sus estudios sobre los daños del borde que podían descubrirse con estereomicroscopios de poca potencia, ignorando que los pulimentos y estriaciones sólo comenzaban a ser visibles con un aumento de 200 diámetros. Muchos de los estudios sobre desgaste de filos estaban prudentemente apoyados en hechos experimentales. Se fabricaron copias modernas de utensilios paleolíticos con un amplio surtido de materiales y se usaron en varios menesteres, a fin de establecer si las huellas de desgaste producidas variaban de unos materiales a

otros. Sin embargo, la mayoría de estos programas constaban de muy pocos experimentos, controlaban escasas variables y tenían un campo de acción demasiado limitado como para obtener resultados aprovechables.

El único programa competente encaminado al análisis de los daños del borde, empleando pocos aumentos, fue llevado a cabo en la Universidad de Harvard por Ruth Tringham y sus discípulos. Cuando se publicaron los resultados de este trabajo en 1974, la principal diferencia comprobable se establecía entre el trabajo hecho sobre materiales "duros" (tales como hueso, cuerno y madera) y el efectuado sobre materiales "blandos" (tales como carne, cuero y vegetales no leñosos). No se encontraron criterios seguros para distinguir entre los diferentes métodos de trabajo, tales como raer, afilar, serrar, cortar, etcétera. También era imposible distinguir entre las señales de deterioro del filo producidas realmente por el uso de un instrumento, por una parte, y, por otra, las pequeñas marcas originadas durante la fabricación o por el roce de los instrumentos contra otros materiales duros a lo largo del dilatado período de enterramiento.

Comencé a investigar por primera vez el microdesgaste en 1972, después de revisar la bibliografía existente sobre el tema y de algunos estudios preliminares. Estos trabajos previos me convencieron de que debía utilizar un orden de aumentos microscópicos más amplio y recurrir a técnicas distintas de las usadas por otros investigadores. Empecé con un programa de experimentos encaminado a proporcionar un sistema de análisis sobre los usos a que habían sido destinados determinados grupos de instrumentos de sílex procedentes de yacimientos ingleses del Paleolítico inferior, es decir, con una antigüedad entre 500.000 y 100.000 años.

Podía disponer de tres microscopios: un estereomicroscopio de reflexión con una escala de aumentos entre seis y cincuenta diámetros; un microscopio óptico con aumentos comprendidos entre 50 y 1000 diámetros, y, mi principal arma de trabajo, un microscopio con luz incidente incorporada y una escala de aumentos entre 24 y 400 diámetros. También utilicé ocasionalmente un microscopio electrónico de barrido, sobre todo para aumentos superiores a los 500 diámetros. Después de hacer copias de instrumentos de piedra del Paleolítico, llevé a cabo una serie de cerca de 200 experimentos, trabajando sobre diversas clases de sustancias comestibles y de otros tipos, empleando varios procedimientos. También sometí ciertos instrumentos a los tipos de desgaste natural que podían o bien producir señales microscópicas análogas a las producidas por el uso humano, o bien borrar estas señales. Dentro de esa línea de trabajo, y gracias a que disponía de un gran número de instrumentos paleolíticos que habían estado sometidos a desgaste por movimientos de tierra, meteorización química y abrasión por materiales fluviales y eólicos, pude comparar este tipo natural de desgaste con mis resultados experimentales.

El descubrimiento esencial obtenido de estos experimentos fue que los pulimentos por microdesgaste sobre los bordes trabajados de las copias modernas se hacían visibles con aumentos comprendidos entre 100 y 400 diámetros, con una iluminación incidente sobre la muestra según un ángulo de 90 grados con el eje óptico del microscopio. Los diferentes tipos de pulimento podían distinguirse fácilmente uno de otro. Encontré que si la actividad había sido serrar o trabajar madera, cortar hueso, partir carne o raer pieles, cada una de ellas producía un tipo característico de pulimento por uso.

Se comprobó que los pulimentos producidos por estos trabajos eran muy estables; no pudieron borrarse de mis copias de instrumentos, ni siquiera con tratamiento químico. Apliqué productos cáusticos que abarcaban las gradaciones completas del *p*H, desde un extremo básico (hidróxido sódico) hasta otro ácido (clorhídrico), sin que se produjese ningún efecto. Sucedió igual con varios disolventes orgánicos. Llegué así a la conclusión de que los pulimentos por uso representaban alteraciones reales y permanentes en la microtopografía del sílex. En consecuencia, parecía evidente que los pulimentos del mismo tipo habían permanecido inalterados en los antiguos artefactos de sílex. Siendo así, se podía deducir de las marcas de microdesgaste observables en un instrumento paleolítico el uso preciso para el que había sido empleado.

Los distintos pulimentos de microdesgaste pueden describirse como sigue.

Pulimentos por uso sobre madera: El borde del instrumento presenta un aspecto similar, prescindiendo de que la madera trabajada haya sido dura, blanda, fresca o seca. El pulimento también es igual con independencia de la manera de usar el artefacto. Se trata de un pulimento "brillante", que refleja un gran porcentaje de la iluminación incidente y con una textura muy lisa. Debido a que el pulimento se desarrolla primeramente sobre las partes salientes de la superficie microtopográfica del sílex, su aspecto general se va afectando en la topografía saliente hasta que la zona de contacto llega a estar completamente pulimentada. Así, si la topografía original del sílex es tosca, la distribución inicial del pulimento tendrá un aspecto reticular. Si el sílex es de grano fino, el pulimento se extiende pronto por toda la superficie. Prescindiendo de la distribución, el pulimento tiene un brillo constante y un aspecto terso.

Pulimentos producidos al trabajar hueso: El borde del instrumento es brillante, pero el pulimento tiene una textura áspera y desigual en vez de la característica tersura del pulimento por madera. Un rasgo distintivo de la tosca textura del pulimento por hueso es la presencia de numerosos hoyuelos sobre la brillante superficie. El pulimento por hueso tiene un desarrollo más lento que el pulimento por madera. Sobre una copia moderna, incluso después de un uso prolongado, el pulimento rara vez tiene un desarrollo muy amplio. Mis experimentos revelaron que no existían diferencias apreciables entre los pulimentos originados en utensilios usados para trabajar huesos cocinados y los originados en aquellos que se habían usado para trabajar huesos crudos; tampoco había diferencias entre los pulimentos de instrumentos usados para trabajar huesos pertenecientes a distintas especies de animales.

Pulimentos por uso sobre piel: En este caso, los bordes del instrumento no desarrollan un único tipo peculiar de pulimento, sino que cambian según el material que ha sido trabajado. Varían desde un pulimento relativamente brillante y de aspecto untuoso (producido al trabajar piel fresca) hasta un pulimento mate apagado (originado al trabajar cuero o piel seca). Estas diferencias son atribuibles a la distinta cantidad de grasa existente en la piel del animal en las diferentes fases. Una piel fresca produce, de forma gradual, un pulimento que no difiere del que se origina al partir carne. Cuando la piel se va secando, disminuye progresivamente su contenido en grasa, con lo que el pulimento del artefacto no sólo se desarrolla más de prisa, sino que también tiene una apariencia más apagada y menos untuosa. Si la piel está completamente seca o curtida, el pulimento toma rápidamente un aspecto apagado y una textura definidamente mate. Prescindiendo de estas diferencias en el pulimento, todos los instrumentos usados para trabajar piel muestran dos tipos

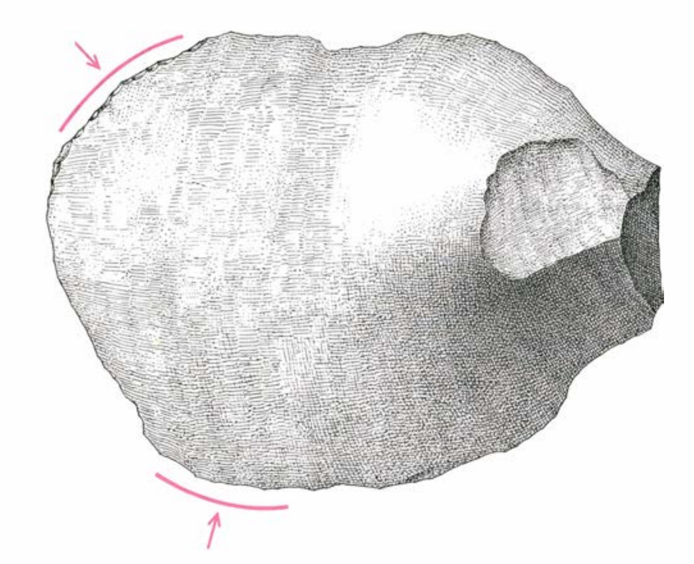
característicos de microdesgaste. Uno es un desgaste relativamente intenso del filo trabajado del instrumento, es decir, una eliminación del sílex que no está causada por rotura o raspadura. Este desgaste da al artefacto lítico un borde marcadamente redondeado. El otro consiste en el desarrollo de trazos superficiales y difusos, formando una superficie longitudinal, que es paralela a la dirección en que se mueve el instrumento. Estas marcas longitudinales difusas se parecen a las estriaciones causadas por otros tipos de materiales, pero no pueden confundirse porque estas últimas resultan mucho más prominentes.

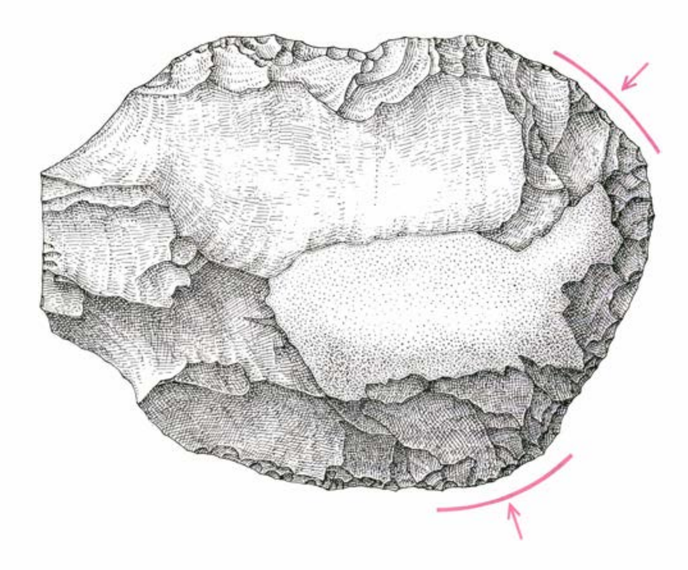
Pulimentos originados al cortar carne: El borde del instrumento que se usa para partir carne y otros tejidos animales blandos desarrolla un pulimento por microdesgaste bastante parecido al pulimento producido al trabajar piel fresca. Este pulimento, sin embargo, se distingue fácilmente de aquellos otros que se producen al trabajar piel seca, hueso, cuerno, madera y vegetales no leñosos. Es un pulimento sumamente untuoso y, al propio tiempo, más apagado que brillante. Así pues, respecto al brillo, el contraste entre este tipo de desgaste y una superficie inalterada de sílex es escaso. Por esta razón, el pulimento originado al partir carne no destaca mucho en una microfotografía, distinguiéndose con claridad en observación visual directa. La característica textura granular del sílex fresco es reemplazada por una textura mate que, aunque parece conservar la microtopografía de la superficie original, tiene en realidad transformadas las elevaciones y depresiones en una superficie semicontinua. Los instrumentos con pulimentos producidos al cortar carne también llevan con frecuencia estriaciones estrechas y cortas.

Pulimentos producidos al trabajar cuerno: Los bordes de los instrumentos usados para trabajar cuerno presentan uno u otro de dos tipos característicos de pulimentos. La diferencia depende de cómo se haya usado el instrumento. Si se ha empleado para raspar, cepillar o grabar cuerno se produce un pulimento muy brillante y liso. Serrando cuerno, se da lugar a un pulimento semejante al que se origina trabajando hueso, o sea, brillante y picado. En el primer tipo, al principio suele desarrollarse un pulimento prácticamente indistinguible del que se produce al trabajar madera. En un desarrollo posterior, la superficie pulimentada presenta pequeñas depresiones esparcidas, dándole una apariencia crateriforme, que es muy diferente tanto de un pulimento por madera como del picado superficial más fuerte, característico del segundo tipo. Mis experimentos con cuerno se llevaron a cabo casi en su totalidad con muestras previamente remojadas en agua durante uno o dos días. El cuerno seco es tan duro que las herramientas de piedra usadas para trabajarlo se vuelven romas y se les estropea el filo antes de conseguir hacer algo. El cuerno remojado es, sin embargo, muy fácil de trabajar.

Pulimentos producidos por plantas no leñosas: El filo de los instrumentos usados para cortar tallos de plantas no leñosas, tales como forraje o helechos adquieren un "brillo de maíz". El rasgo característico es una superficie muy lisa altamente reflectante, de apariencia "fluida". Si existen estriaciones, aparecen generalmente como "rellenas". Al mismo tiempo, las superficies pulimentadas del filo trabajado desarrollaron unos curiosos hoyitos en forma de cometa. Como indica la expresión, el brillo de maíz se encuentra la mayor parte de las veces sobre sílex falciformes usados por los agricultores neolíticos para recolectar algunas especies cultivadas de la familia de las gramíneas. Sin embargo, según pude descubrir, también se habían cortado plantas no leñosas en el Paleolítico inferior, y los instrumentos empleados habían sufrido el mismo tipo de pulimento.

Los pulimentos por trabajo sólo capacitan al investigador para deducir qué materiales fueron tratados con los diferentes instrumentos de sílex. Sin embargo, si se quiere determinar cómo se usaron los instrumentos, hay que basarse en otros tipos de pruebas de desgaste microscópico. Quizás el más importante sea la distribución y orientación de marcas de desgaste longitudinales del tipo de las estriaciones. Otras clases de pruebas comprenden la situación y naturaleza de los daños sufridos por el borde, así como la posición y extensión de las porciones que presentan pulimentos por trabajo. Todas las pruebas de este tipo deben considerarse en relación con el tamaño y forma general del instrumento. En términos generales, una vez que se ha hecho un inventario de los diferentes tipos de microdesgaste que pueden verse en un determinado artefacto, hay que plantearse el modo en que ha debido





INSTRUMENTO EN LASCA del Paleolítico; este "raspador lateral" procedente de Hoxne –un yacimiento inglés achelense– estaba entre unos 800 instrumentos de sílex examinados por el autor en su búsqueda de pruebas de microdesgaste. La otra cara de la lasca *(derecha)* muestra todavía algunas de las superficies *(zona clara)* correspondientes a la parte externa del nódulo de sílex del que se separó la lasca. Las líneas y flechas *(en color)* indican los bordes trabajados del instrumento. Los prehistoriadores habían supuesto que los raspadores se utilizaban para manipular pieles de animales. Las huellas o marcas de microdesgaste que se han podido descubrir en la lasca corroboran esta suposición.





COPIA Y ORIGINAL, comparadas en estas microfotografías. El borde del instrumento que se ve a la izquierda fue usado por el autor para raspar pieles secas de cerdo durante una hora. El borde, visto con un aumento de 300 diámetros, desarrolla un pulimento por trabajo mate apagado, y comienza a redondearse por desgaste. A la derecha puede verse, con el mismo aumento, el borde del raspador de Hoxne, dibujado en la ilustración de la página anterior. Presenta el mismo pulimento mate apagado así como el redondeamiento por desgaste formado sobre la copia.

utilizarse para adquirir los rasgos observados. Por ejemplo, la presencia de huellas de microdesgaste en los dos lados de un borde trabajado, junto con estriaciones paralelas a ese borde, constituyen las pruebas más evidentes de que se empleó el instrumento para serrar o cortar. El análisis del pulimento por trabajo podría indicar entonces qué material fue serrado o cortado.

Después de establecer seis categorías claras de pulimentos, estuve en condiciones de aplicar mis resultados experimentales a una selección de artefactos paleolíticos. Sin embargo, un escéptico colega sugirió que primero sometiese mi técnica analítica a una prueba secreta. Este colega, Mark H. Newcomer, de la Universidad de Londres, tenía serias dudas sobre la validez del análisis de microdesgaste. Acordamos que él fabricaría varias copias de instrumentos de sílex y que trabajaría con ellas sobre diversos materiales. Después de dejar constancia del uso dado a los instrumentos y haberlos limpiado, me los enviaría para proceder a su análisis. Luego, nos reuniríamos para comparar mis deducciones con su informe del uso real dado a los instrumentos. Newcomer hizo quince copias de instrumentos antiguos de sílex y llevó a cabo diversas clases de trabajos usando un total de 16 folios de instrumentos.

Los resultados de esta prueba secreta fueron instructivos. Para que fuese más fiable, el número de artefactos era pequeño. Sin embargo, identifiqué las porciones trabajadas de los bordes de instrumentos en 14 casos sobre 16. Para 12 de los bordes pude reconstruir la forma en que se había usado el instrumento y, para otros 10, deduje el tipo de material sobre el que se había trabajado.

Algunas de las deducciones estaban notablemente ligadas a las marcas. Por ejemplo, Newcomer había desollado una liebre con un instrumento de dos bordes, usando uno de ellos para el despellejamiento propiamente dicho y el otro para cortar las partes de las extreminades que habían conservado la piel durante la manipulación secreta. Identifiqué el desgaste sobre el filo usado para desollar como un pulimento originado al cortar carne. (No tenía forma de saber que en este caso la carne estaba a menos de un milímetro por debajo de la piel.) El microdesgaste del borde opuesto lo interpreté como el resultado de partir una articulación.

Newcomer había cortado con otro instrumento carne fresca colocada sobre una tabla de madera de trinchar. Pude distinguir entre el microdesgaste causado al cortar la carne y el desgaste incidental originado por el contacto entre el sílex y la tabla de trinchar. Más todavía, algunas de mis interpretaciones equivocadas no carecían de justificación. Por ejemplo, Newcomer había utilizado el borde de un instrumento de sílex para cortar carne congelada, lo cual produce pocas marcas de desgaste. Sin embargo, había cortado la carne sobre una tabla de madera, y el contacto con ésta había originado señales identificables. Interpreté el microdesgaste resultante como característico de un instrumento usado muy suavemente sobre madera. Como las pruebas de Newcomer constituyeron la primera comprobación de la validez del análisis de microdesgaste de alta resolución, encontré los resultados muy alentadores.

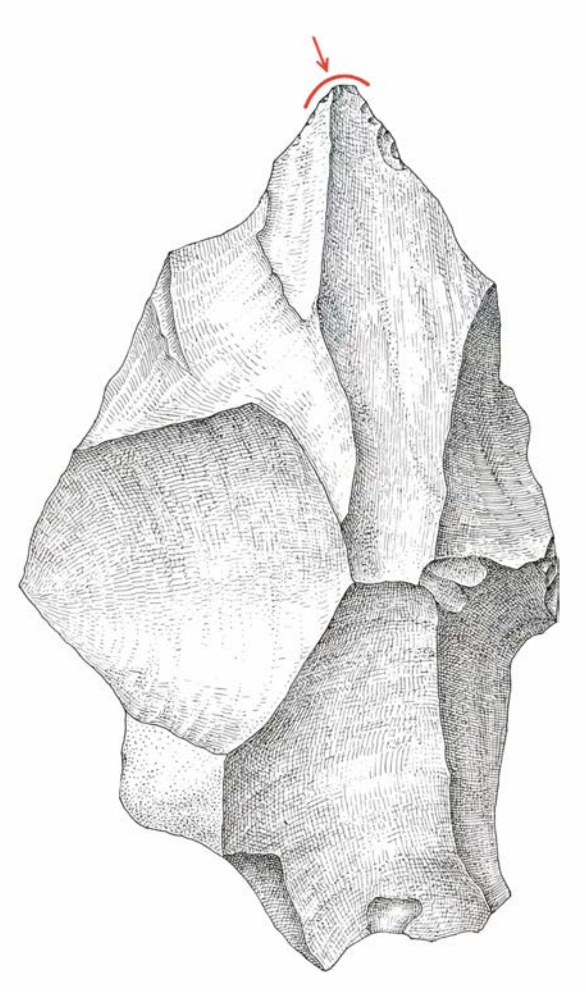
Había llegado el momento de aplicar esta técnica a la selección de artefactos paleolíticos. Había tres yacimientos clásicos británicos del Paleolítico inferior que reunían las condiciones deseadas. En primer lugar, los instrumentos de sílex de los tres yacimientos se hallaban bien conservados; no estaban recubiertos por una capa superficial que pudiese ocultar o destruir las pruebas de microdesgaste y, en general, no habían sufrido daño por abrasión natural. En segundo lugar, todos los artefactos habían sido excavados recientemente, cuidando que su posición estratigráfica hubiese sido estrictamente controlada y que su manipulación y almacenamiento se hubiese hecho cuidadosamente a fin de eliminar el riesgo de un daño posterior a su excavación. Los yacimientos eran: el "Golf Course", en Clancton-on-Sea, condado de Essex; el "Lower Loam", en Swanscombe, condado de Kent; y el "Lower Industry" (principalmente), en Hoxne, condado de Suffolk.

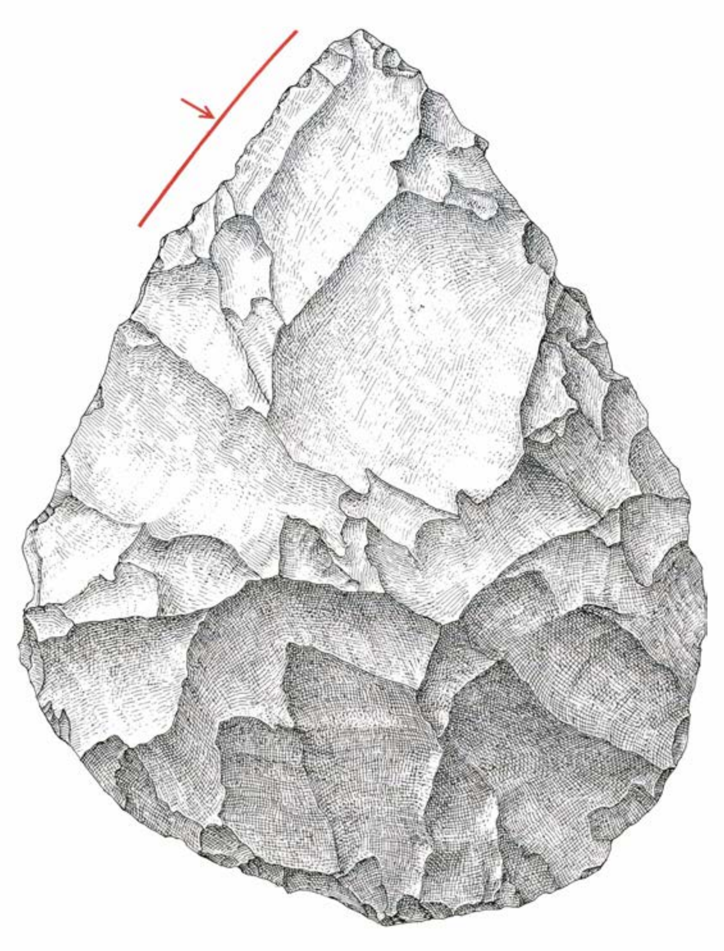
El yacimiento de Clancton ha dado el nombre de clanctoniense a un tipo completo de industria lítica en sílex que floreció hace unos 250.000 años, durante las primeras fases del período interglacial Mindel-Riss. Los característicos instrumentos en lascas de la industria clanctoniense se fabricaron por percusión, sacando primero lascas brutas de nódulos de sílex y desbastando alguna de ellas

hasta darles la forma deseada. Algunos de los núcleos de sílex menos aprovechables se emplearon también como instrumentos.

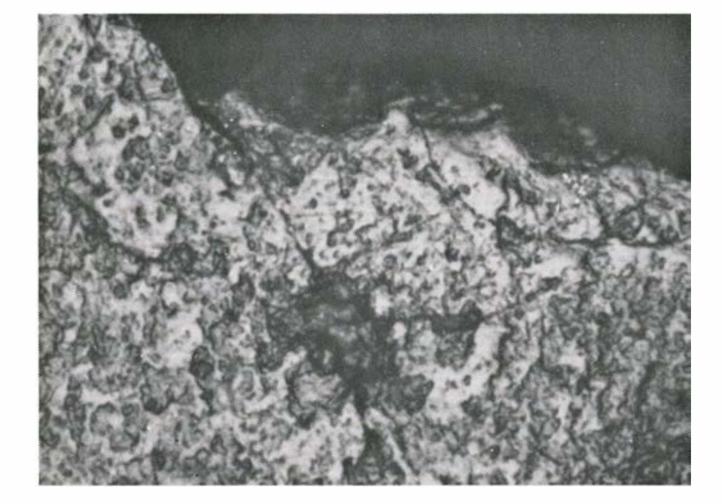
Lower Loam, en Swanscombe, es un yacimiento algo posterior al clanctoniense, habitado durante el mismo período interglacial Mindel-Riss. Los artefactos líticos procedentes de Lower Loam abarcan instrumentos hechos con lascas. Hoxne, yacimiento más tardío, ha proporcionado refinados artefactos que están trabajados por ambos lados. Estos instrumentos bifaciales se atribuyen tipológicamente al achelense, una industria del Paleolítico inferior cuyo nombre procede del yacimiento de Saint-Acheul, en Francia, donde se descubrieron por primera vez. Los estratos de Hoxne también contienen abundantes lascas de sílex, muchas de las cuales son restos no aprovechables de la fabricación de herramientas bifaciales.

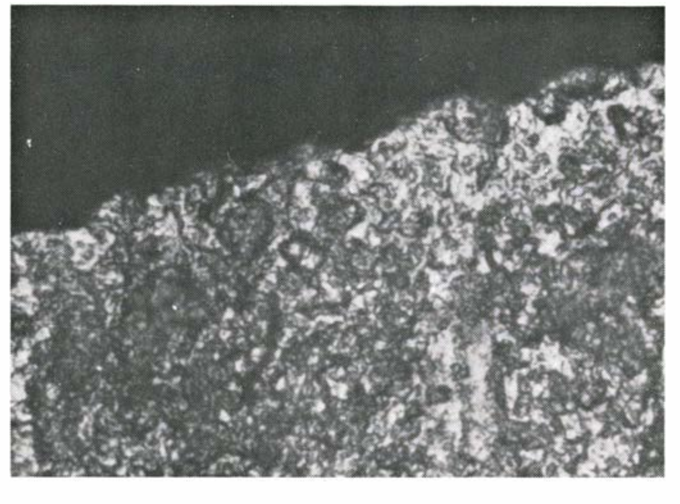
El número total de instrumentos en condiciones apropiadas para el análisis de microdesgaste no era grande. El grupo de Clancton comprendía 144 artefac-





HERRAMIENTAS BIFACIALES procedentes de dos yacimientos del Paleolítico inferior de Inglaterra. A la izquierda, un "chopper" hallado en Clancton-on-Sea, y junto a él un "hacha de mano" procedente del yacimiento achelense de Hoxne. Las líneas y flechas *(en color)* muestran sus filos trabajados. El análisis de microdesgaste de los mismos nos pone claramente de manifiesto que sirvieron para funciones diferentes.





DOS FILOS TRABAJADOS aumentados 300 veces. A la izquierda, el borde de "taladro" del bifaz de Clancton. La presencia de pulimentos originados por madera indican que el supuesto "chopper" se utilizó para trabajar este material; el desgaste sobre la punta de taladro, no visible aquí, indica que el instrumento se usó para hacer orificios. El "hacha" de Hoxne *(derecha)* presenta el pulimento por trabajo de cortar carne.

tos procedentes de un nivel de grava del yacimiento de Golf Course y otros 102 de un nivel de marga. Algunas de las lascas podían acoplarse al núcleo primitivo del que habían sido separadas, lo que ponía de manifiesto que la fabricación había tenido lugar allí mismo. Al deducir, mediante las pruebas de microdesgaste, que los instrumentos en lascas se habían usado en una misma época para trabajar carne, madera, piel y, en ocasiones, hueso, se podía sugerir que el yacimiento de Golf Course había permanecido ocupado probablemente por algún tiempo, en vez de haber sido un esporádico campamento de cazadores. Las actividades fundamentales en el yacimiento fueron la manipulación de madera y carne.

De los artefactos de Clancton que yo examiné, 22 eran herramientas bifaciales toscas clasificadas tradicionalmente como *choppers*. El microdesgaste demostró que sólo dos de los 22 se habían usado realmente como herramientas. Esto supone una utilización del orden del 10 por ciento, la mitad de lo calculado para las lascas encontradas en el sitio, aproximadamente. Estas proporciones relativas sugieren que los fabricantes de utensilios de Clancton preferían en primer lugar utilizar sus núcleos de sílex para producir lascas, en vez de bifaces.

Había sesenta y seis instrumentos en lascas procedentes de Lower Loam, en Swanscombe, en condiciones apropiadas para el análisis de microdesgaste, de los que solamente cuatro presentaban verdaderamente señales de uso. Las cuatro lascas mostraban rasgos de microdesgaste parecidísimos a los que se veían en los instrumentos de Clancton, aunque la muestra de Swanscombe era demasiado escasa para permitir sacar conclusiones de esta coincidencia.

Los artefactos de Hoxne comprendían un grupo de instrumentos (Lower Industry; nivel 3, oeste) con escaso o ningún daño por abrasión. Estudié el conjunto completo procedente de este nivel, numerando 408 instrumentos. También analicé una muestra de artefactos tomados al azar procedentes de otros estratos de los yacimientos de Lower Industry y Upper Industry. La industria achelense en Hoxne, con su intensidad en la manufactura de herramientas bifaciales, se caracteriza por la gran cantidad de lascas que deben considerarse como instrumentos potenciales, a pesar de que la mayoría son seguramente desechos de fabricación de instrumentos con el borde demasiado delgado y frágil para hacer con ellos instrumentos en lascas. Verdaderamente, el análisis de microdesgaste reveló que solamente el 9 por ciento de las lascas procedentes de todos los niveles del yacimiento de Lower Industry, presentaban realmente pruebas de desgaste. Los artífices de los instrumentos de Hoxne los usaron para partir carne, trabajar madera y piel y para taladrar madera y hueso. También es interesante que algunos fueron empleados para trocear o partir plantas no leñosas. Es posible que tales cazadores recolectasen cañas y grandes helechos para hacer yacijas. La manipulación de la carne no se hacía exclusivamente con instrumentos de lascas: dos de las "hachas de mano" de Lower Industry muestran el pulimento característico de los instrumentos usados para cortar carne.

Entre los instrumentos del yacimiento Upper Industry, en Hoxne, había un pequeño número de instrumentos en lascas de los llamados tradicionalmente "rascadores laterales" cuyo uso se ha supuesto ligado al curtido de pieles. El microdesgaste de estos instrumentos apoya esta suposición tradicional: la mayoría de los rascadores laterales presentan el pulimento característico de los instrumentos usados para trabajar piel.

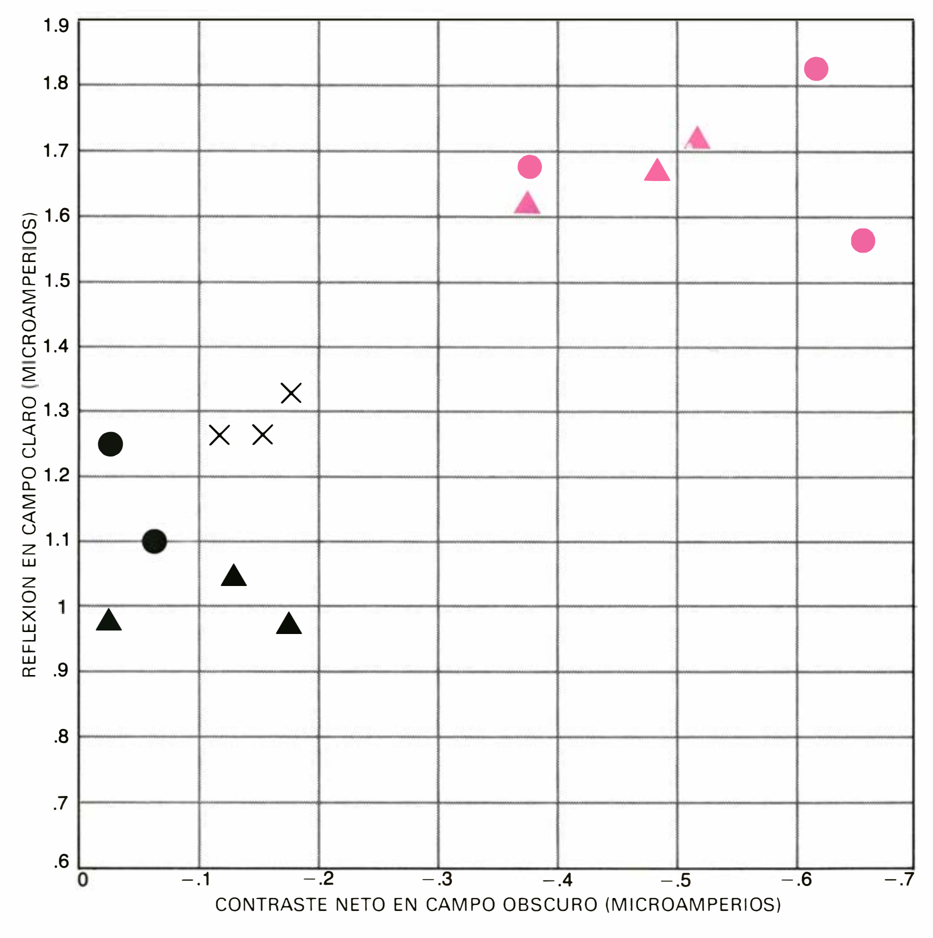
Por citar un nuevo ejemplo de análisis de microdesgaste, se encontró que un instrumento bifacial de Clancton presentaba, en su superficie trabajada, pulimento por uso sobre madera. Un examen posterior reveló un desgaste por utilización que sólo podía provenir de un movimiento rotatorio como el que se produce en una perforación; se había girado el instrumento en el sentido de las agujas del reloj, al mismo tiempo que se le iba presionando hacia abajo. Unos tipos análogos de desgaste habían aparecido en los instrumentos en lascas usados para taladrar. El esquema de las huellas indicaba que los trabajadores de madera de Clancton de hace unos 200.000 años eran consecuentemente diestros.

El supuesto desperdicio representado por el orden de utilización del 9 por ciento de lascas en Hoxne puede ser más aparente que real; la mayoría de las lascas eran desechos de la fabricación de bifaces, inadecuados para usarse como instrumentos. La abundante utilización de sílex en Clancton no puede explicarse igualmente en otros lugares como desperdicios de la producción de instrumentos bifaciales. Quizá la gran cantidad de desechos en ambos yacimientos se explique mejor porque la caliza dura, una excelente materia prima para la fabricación de herramientas de piedra, pue-

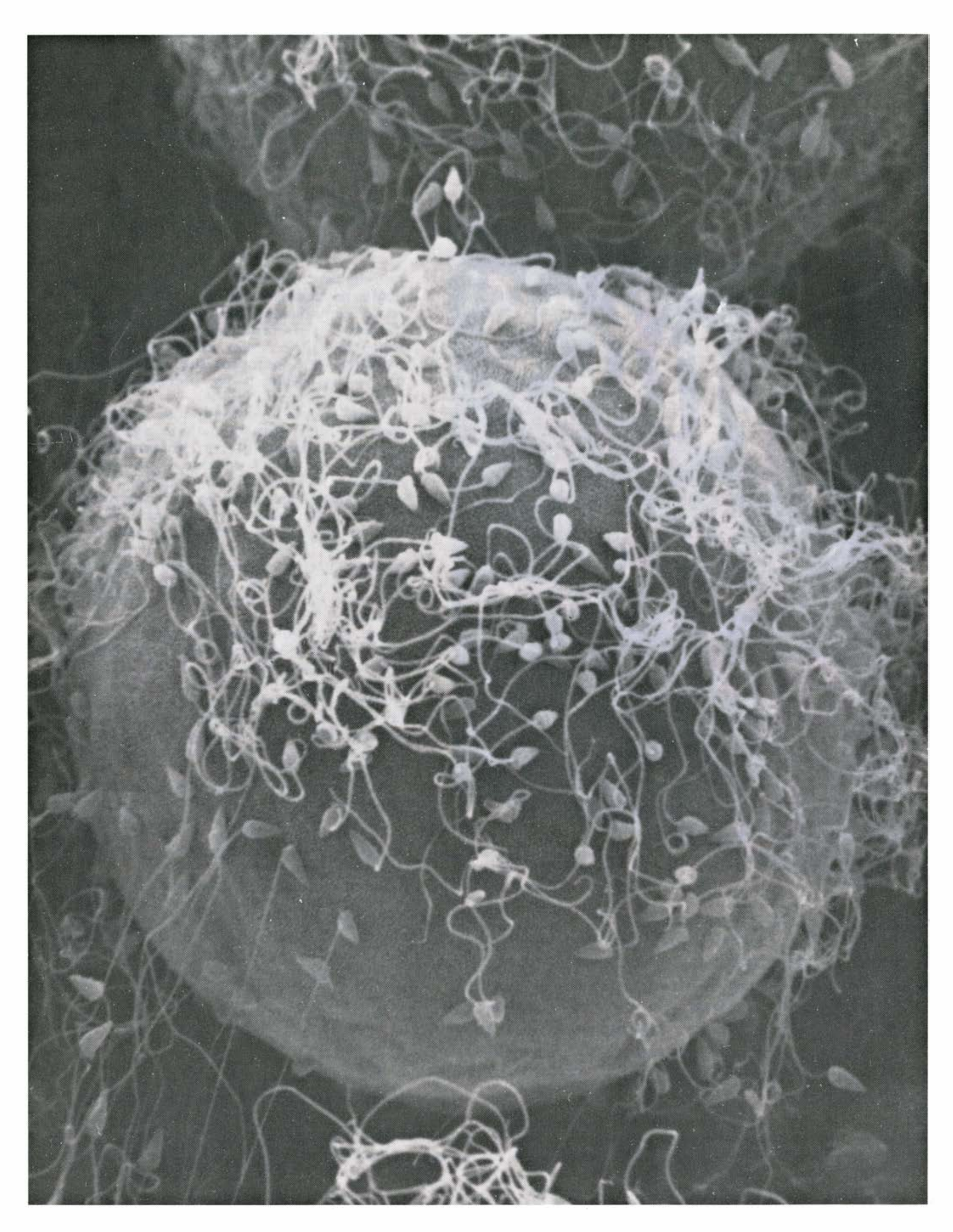
de encontrarse fácilmente casi en todas las partes del sureste de Inglaterra: en los aluviones fluviales, las playas y en otros depósitos superficiales, y, por supuesto, en los propios afloramientos calizos.

El análisis de microdesgaste de los pulimentos por trabajo en este grupo de instrumentos del Paleolítico inferior proporcionó la primera prueba directa e inequívoca de los tipos de actividad humana que tuvieron lugar en asentamientos ingleses de hace unos 250.000 años. Estos hallazgos han puesto de manifiesto que por fin ha alcanzado su mayoría de edad un nuevo y prometedor método de investigación arqueológica. Ahora es posible, suponiendo que los instrumentos se hayan conservado convenientemente, determinar en la mayoría de los casos, no solamente cómo se utilizaron los antiguos artefactos de sílex, sino también qué destino tenían.

Parece muy probable, aunque no está todavía demostrado que el microdesgaste presente en instrumentos fabricados en materiales pétreos distintos del sílex –tales como obsidiana, calcedonia, e incluso basaltos y cuarcitas de grano fino– puede interpretarse de un modo análogo. Yo lo he comprobado en una calcedonia de grano fino, del Africa austral; los experimentos demuestran que el material conserva pulimentos de microdesgaste directamente comparables con los que se encuentran en una caliza dura. La información procedente de futuros estudios de microdesgaste deberá permitir a los prehistoriadores discutir con seguridad cada vez mayor la tecnología y economía del hombre primitivo.



BRILLO RELATIVO de varios pulimentos por trabajo en función de dos parámetros, esquematizados en esta gráfica de reflectividad. Los valores de las ordenadas indican la cantidad de luz reflejada procedente de una zona patrón de la superficie pulimentada, bajo iluminación normal en campo claro, medida en microamperios mediante un fotómetro. Los valores en abscisas indican las diferencias entre la reflectividad de las superficies pulimentadas y no pulimentadas de un instrumento bajo iluminación en campo obscuro: cuanto más pequeña es la diferencia, más áspera es la textura de la superficie pulimentada. El más brillante y liso de los pulimentos por trabajo resultó ser el "brillo de maíz" producido al cortar los tallos de plantas no leñosas ricas en partículas de ópalo. Las figuras en color indican las lecturas tomadas en hojas de segar de sílex del Neolítico, procedentes de Siria. Los triángulos coloreados indican las lecturas en hojas neolíticas de Baviera. Los pulimentos por trabajo formados sobre copias modernas al trabajar pieles grasientas *(triángulos negros)* estaban entre los más apagados y ásperos; los pulimentos producidos al trabajar pieles secas *(topos negros)* eran más ásperos y, al propio tiempo, más brillantes. Los pulimentos producidos al trabajar madera *(cruces negras)* resultaban más lisos y brillantes.



MILES DE ESPERMATOZOOS, adheridos a la superficie de un óvulo de erizo de mar, aparecen en esta microfotografía, obtenida por Mia Tegner. La superficie registrada aquí no es la membrana plasmática (la membrana externa del óvulo en sí) sino una cápsula denominada capa vitelina, que contiene los receptores de los espermatozoos. Pese al elevado número de espermatozoos, sólo uno de ellos logrará fecundar el óvulo. Si penetrara más de uno (fenómeno conocido como polispermia) se produciría un exceso de cromosomas paternos, y el desarrollo del embrión quedaría abortado. Los óvulos han desarrollado ciertos mecanismos que evitan la polispermia. Aumento: 2880 diámetros

El proceso de la fecundación

La unión de un espermatozoo y un óvulo desencadena una serie de cambios transitorios en la concentración de iones, que impiden la adición de más espermatozoos e inician el desarrollo embrionario

David Epel

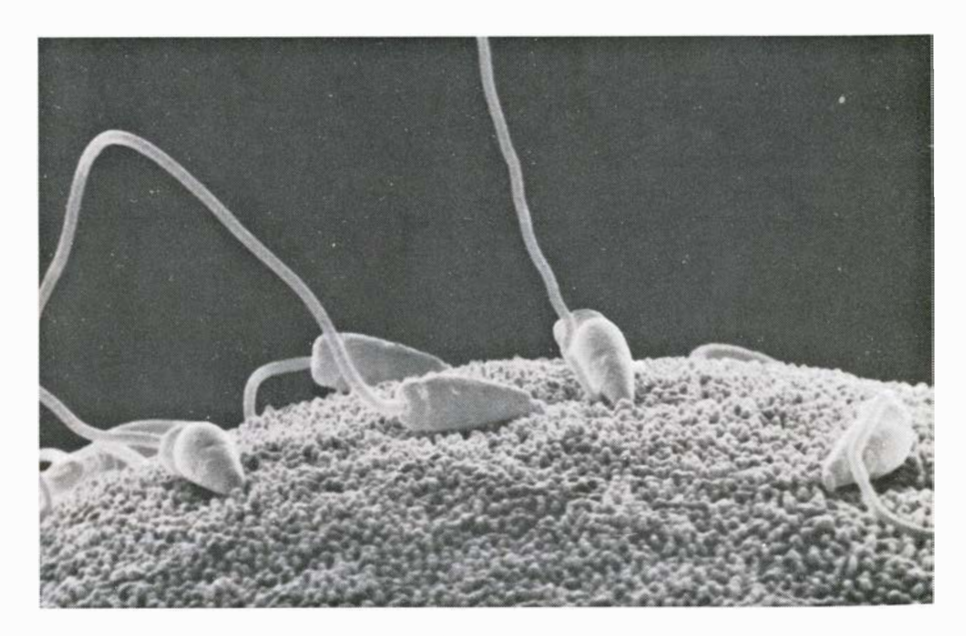
La interacción entre un espermatozoo y un óvulo marca una línea divisoria entre la vida y la muerte. Si ambas células interactúan con éxito, y tiene lugar la fecundación, sus respectivos núcleos, que contienen cada uno la mitad de la dotación completa de cromosomas, se unen, y se inicia el desarrollo de un nuevo individuo. Si la interacción no es satisfactoria, las células mueren al cabo de unas pocas horas, o, a lo sumo, a los pocos días. Debido a ello, los organismos vivos emplean una considerable cantidad de energía, tanto fisiológica como de comportamiento, en asegurar que estas dos células lleguen a encontrarse y se produzca la fecundación.

El proceso de la fecundación puede dividirse en tres fases principales: el reconocimiento del óvulo por el espermatozoo, la regulación de la entrada de espermatozoos en el óvulo, de forma que sólo el material genético de un espermatozoo se una al del óvulo, y la activación del metabolismo del óvulo, hasta entonces aletargado, para que dé comienzo la división celular y el desarrollo embrionario. La intensa investigación bioquímica de estos últimos años ha empezado ya a proporcionar una descripción de estos acontecimientos a nivel molecular. Las ideas surgidas a partir de estos estudios pueden ayudarnos a comprender otros tipos de transformaciones celulares, como las que ocurren en el cáncer, y pueden también procurar nuevos enfoques al control de la fertilidad humana.

Los huevos de invertebrados marinos, particularmente los de equinodermos (como el erizo de mar y las estrellas de mar), constituyen el material clásico para el estudio de la fecundación. Estos animales liberan un elevado número de gametos (óvulos o espermatozoos), de los que sólo unos pocos llegan a interactuar en el agua marina, dando una descendencia viable. Los erizos de mar, por ejemplo, tienen una época de cría que dura de 3 a 8 meses, durante la cual una hembra puede liberar hasta 400 millones de óvulos, y un macho hasta 100.000 millones de espermatozoos. Un hecho de notable interés para el investigador es que tanto los óvulos como los espermatozoos se extraen de estos organismos con suma facilidad, pudiendo efectuarse la conjugación *in vitro* en el laboratorio. Para ello, simplemente se suspenden los óvulos en agua marina, y se añade una pequeña cantidad de espermatozoos, agitando. A los pocos segundos los óvulos son fecundados sincrónicamente, lo que permite estudiar los fenómenos que se presentan no sólo en un huevo, sino también en un gran número de ellos.

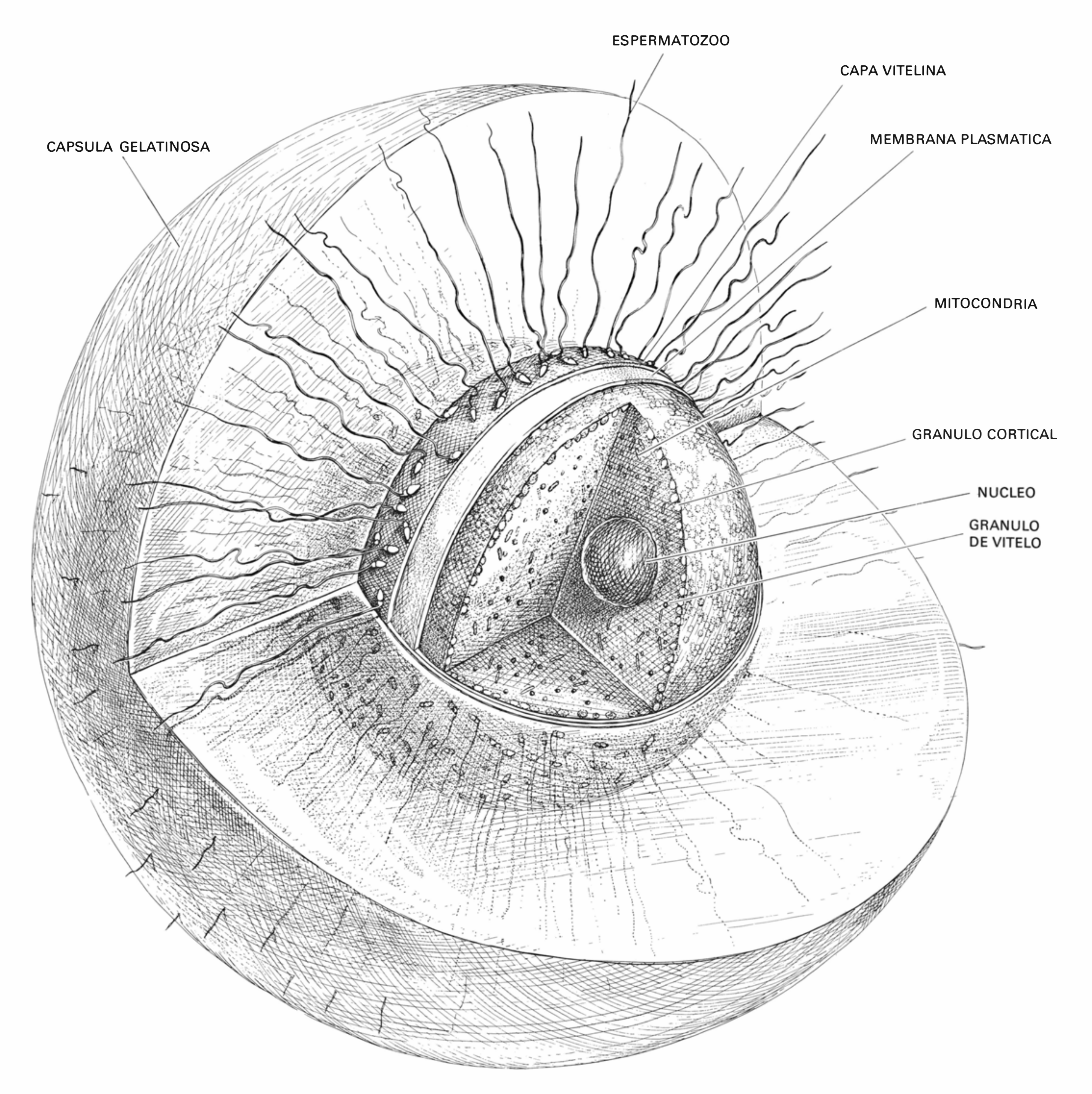
En 1877, el zoólogo suizo Hermann Fol observó a través del microscopio cómo un espermatozoo de estrella de mar se adhería a un óvulo y lo fecundaba; con ello se daba fin a siglos de especulación. Sin embargo, sólo recientemente se ha empezado a hallar la respuesta a algunas de las cuestiones fundamentales de los mecanismos de la fecundación. Uno de los problemas que se han planteado durante largo tiempo los biólogos es cómo el espermatozoo reconoce específicamente al óvulo. Aunque el espermatozoo parece alcanzar al óvulo en razón de su elevado número y por casualidad, ciertos mecanismos deben impedir que se una a cualquier otro tipo de célula con que pueda encontrarse durante el recorrido.

Se supone que el reconocimiento específico del óvulo por el espermatozoo tiene lugar cuando éste toma contacto con la cápsula gelatinosa que envuelve al óvulo. Ciertas sustancias de la cápsula gelatinosa interactúan con la parte delantera de la membrana plasmática del espermatozoo (que es su membrana más externa). Esta región del espermatozoo, denominada acrosoma (del griego: cuerpo



EN LA SUPERFICIE DE UN HUEVO FECUNDADO de erizo de mar se observan varios espermatozoos unidos a la capa vitelina. La superficie está recubierta de diminutas proyecciones de la membrana plasmática, denominadas microvilli. Microfotografía de E. William Byrd.

puntiagudo), queda alterada a consecuencia de la interacción, y libera enzimas digestivos que permiten al espermatozoo perforar una abertura en las capas que envuelven al óvulo, de forma que llegue a alcanzar la superficie del óvulo propiamente dicho. En los invertebrados, como el erizo de mar, el acrosoma sufre simultáneamente un notable cambio estructural: segrega un delgado filamento. Este filamento se llama acrosómico; se engancha a la membrana vitelina del óvulo, situada por debajo de la cápsula gelatinosa. El lugar de fijación es aparentemente una proteína receptora asociada a la capa vitelina que reconoce y enlaza con una proteína complementaria del filamento acrosómico. Kenji Aketa y sus colegas, de la Universidad de Nagoya, así como Victor D. Vacquier y sus colegas, de la Universidad de California en Davis, han aislado independientemente las proteínas complementarias. Estos autores han observado que, preparando anticuerpos contra las proteínas receptoras y añadiéndolos a una suspensión de óvulos antes de la adición



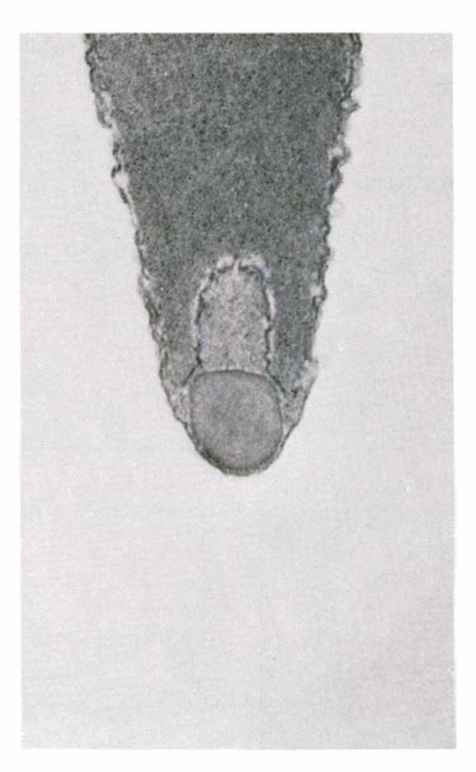
ESQUEMA DE UN OVULO DE ERIZO DE MAR en el instante de la fecundación. El óvulo tiene 75 micras de diámetro y está totalmente recubierto por una gruesa cápsula gelatinosa, a través de la cual penetra el espermatozoo, probablemente perforándola por medio de enzimas digestivos. El espermatozoo se une luego a receptores de la capa vitelina. Inmediatamente por debajo de la membrana plasmática, hay miles de gránulos corticales, de una micra, que se fusionan con la membrana plasmática segundos después de la fecundación y liberan su contenido en el espacio que queda entre la membrana y la capa vitelina. El protoplasma del óvulo contiene numerosos gránulos de vitelo y mitocondrias (que proporcionan respectivamente alimento y energía química) y un único núcleo, que posee la mitad de la dotación normal de cromosomas. Esto último es el resultado de la división reductora, llamada meiosis, que tiene lugar durante una fase avanzada de la maduración del óvulo. La unión de un espermatozoo con el óvulo proporciona la otra mitad de la dotación cromosómica normal, y activa la embriogénesis.

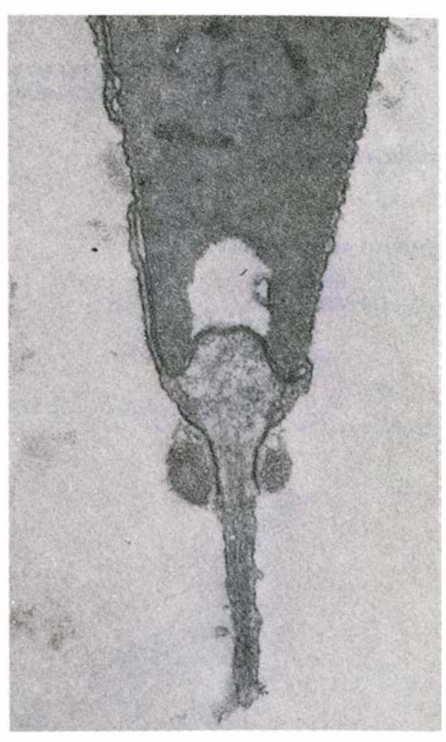
de espermatozoos, no se produce la fecundación.

Una vez que la punta del filamento acrosómico ha interactuado con el receptor, perfora la membrana vitelina y se fusiona con la membrana plasmática del óvulo, situada inmediatamente por debajo. Las membranas plasmáticas de ambas células se unen formando un puente que progresivamente va aumentando de tamaño, hasta que el espermatozoo queda completamente rodeado por la membrana plasmática del óvulo, quedando incorporado en el interior del óvulo.

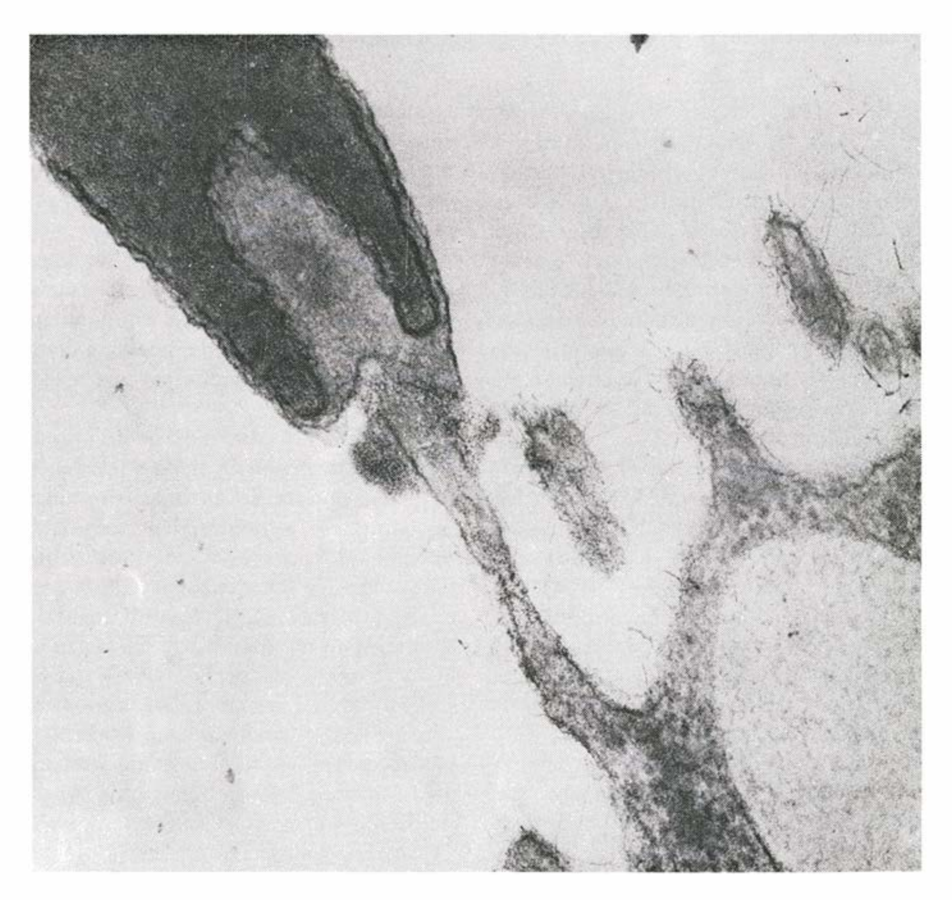
Aunque es un único espermatozoo el que normalmente se fusiona con la membrana plasmática del óvulo y penetra en él, muchos otros espermatozoos se adhieren también a la superficie del óvulo durante la inseminación. Por medio de fotografías efectuadas con microscopio electrónico de barrido por Mia Tegner en mi laboratorio de la Scripps Institution of Oceanography, se ha observado que, en condiciones de saturación, pueden unirse a un único óvulo hasta 1500 espermatozoos. Aunque esta superabundancia es necesaria para asegurar que al menos un espermatozoo llegue a fecundar el óvulo, puede resultar, en potencia, causa de problemas: si más de un espermatozoo penetra en el óvulo, fenómeno que se conoce como polispermia, el número de cromosomas será mayor que el de una dotación normal, y el desarrollo se detendrá en los primeros estadios de la embriogénesis. Por ello, las especies animales han debido desarrollar mecanismos que impiden que más de un espermatozoo penetre en el óvulo.

Los trabajos realizados en los años 50 por Lord Rothschild y Michael M. Swann en Inglaterra, con óvulos de erizo de mar, sugieren la existencia de dos barreras distintas que evitan la polispermia: un bloqueo rápido e incompleto, que tiene lugar durante los primeros segundos del contacto espermatozoo-óvulo, seguido de otro bloqueo más lento, pero también más completo. La naturaleza del primero de ellos ha sido descubierta recientemente por Laurinda Jaffe durante los experimentos realizados para preparar su tesis doctoral en la Universidad de California en Los Angeles. Introdujo microelectrodos en un óvulo de erizo de mar y midió los cambios de tensión entre el interior y el exterior de la célula, inducidos por la fecundación. Observó que, aproximadamente un segundo después del anclaje del espermatozoo, existía un flujo de iones sodio hacia el interior de la célula, provocando una breve variación en el potencial, pa-

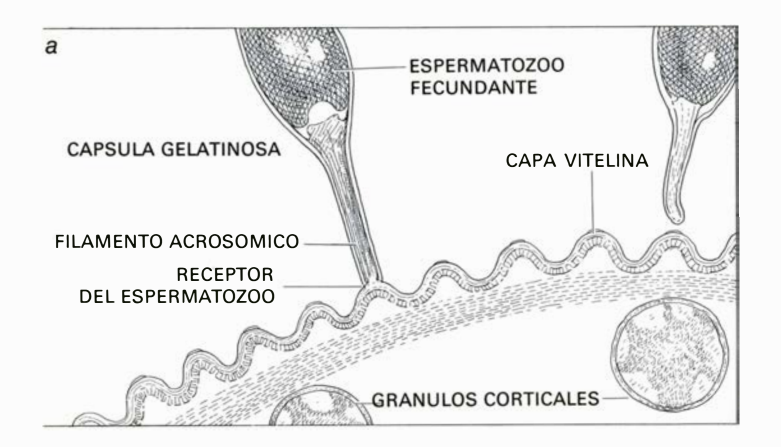


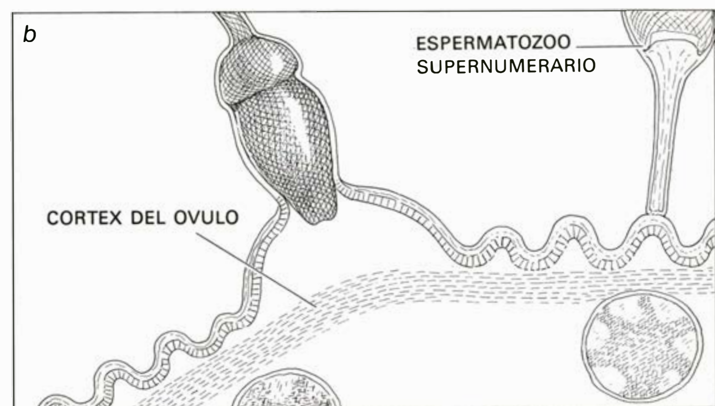


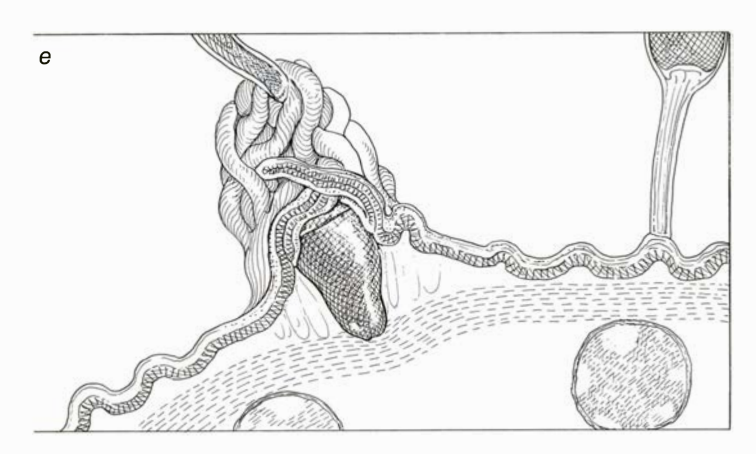
IMPORTANTES CAMBIOS ESTRUCTURALES, después de la interacción inicial entre el espermatozoo y el óvulo, que ocurren en los espermatozoos de los erizos de mar y de otros invertebrados. La parte delantera de la cabeza del espermatozoo segrega un delgado filamento denominado filamento acrosómico que penetra en la capa vitelina del óvulo y se fusiona con la membrana plasmática. Las dos fotografías, obtenidas con un microscopio electrónico de transmisión por Frank Collins, muestran un espermatozoo del erizo de mar *Lytechinus pictus* antes y después de la formación del filamento acrosómico. El aumento es de 45.500 diámetros.

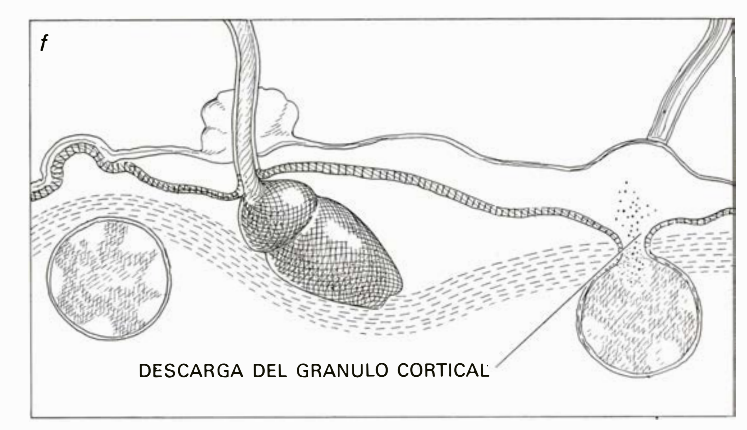


MOMENTO DE LA FUSION de la punta del filamento acrosómico *(izquierda)* con un microvilli del óvulo. Mediante esta fusión se crea un puente citoplasmático por el que el espermatozoo penetra en el óvulo. La fotografía de microscopio electrónico de transmisión fue obtenida por Collins.









LA REACCION CORTICAL se induce por la fusión del espermatozoo fecundante con el óvulo. En esta secuencia, el filamento acrosómico del espermatozoo penetra en la capa vitelina y se une a la membrana plasmática del óvulo *(a, b)*. **Los microvilli próximos al espermatozoo se prolongan y entrelazan por encima de la cabeza del espermatozoo, arrastrándolo al interior del óvulo** *(c-e)*. **La entrada del espermatozoo va acompañada de la fusión de 15.000 gránulos corticales con la membrana plasmática del óvulo y la liberación de su contenido al espacio**

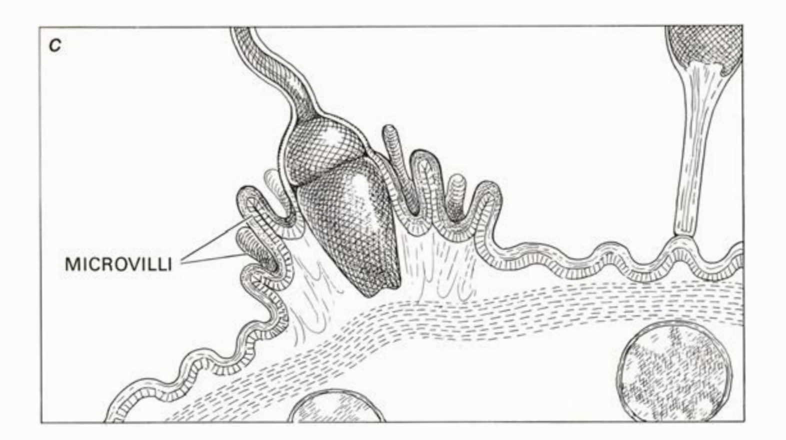
recida a la del impulso nervioso. Al parecer, este cambio de tensión es lo que impide la entrada en el huevo de espermatozoos supernumerarios. Como confirmación de su teoría, Jaffe obtuvo que, al incrementar artificialmente la tensión a través de la membrana de óvulos hasta un nivel observado normalmente después de la fecundación (para lo que utilizaba un "fijador de tensión") y añadiendo entonces espermatozoos, no se producía fecundación alguna.

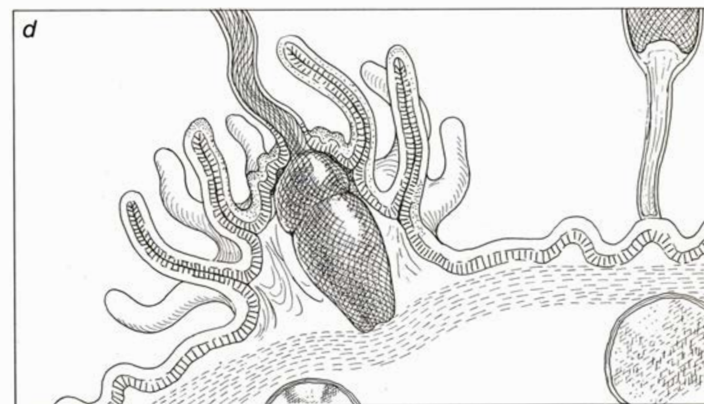
Las investigaciones sobre el segundo bloqueo de la polispermia se han dirigido hacia la reacción cortical: un masivo cambio estructural que se produce en el huevo poco después de la fecundación. Como hemos visto, la membrana plasmática del óvulo no fecundado está rodeada por la delgada membrana vitelina. Inmediatamente por debajo de la membrana plasmática se halla la capa cortical, que contiene alrededor de unas 15.000 pequeñas vesículas llamadas gránulos corticales. Cada uno de éstos tiene un diámetro de aproximadamente una micra (10^{-3} milímetros). Unos 25 a 35 segundos después de que el filamento del acrosoma del espermatozoo penetre en la membrana vitelina, los gránulos corticales se fusionan con la membrana plasmática y descargan su contenido en el espacio que queda entre la membrana plasmática y la vitelina, situada por encima de aquélla. Este proceso se inicia en el lugar en que se unen el óvulo y el espermatozoo, propagándose rápidamente por toda la superficie del huevo, y completándose veinte segundos más tarde.

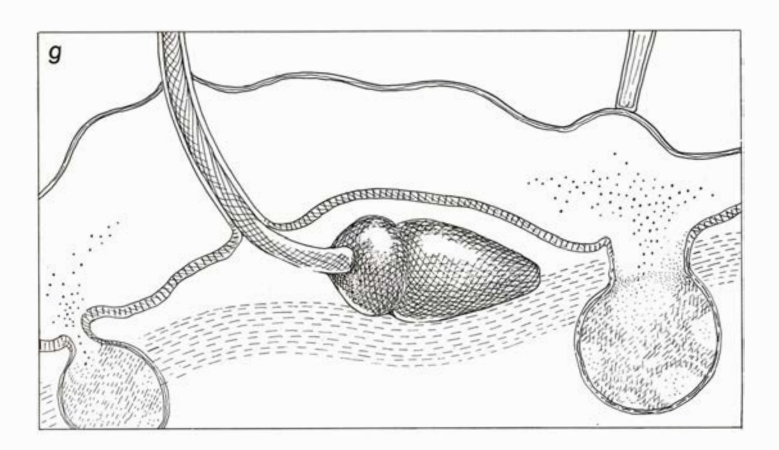
El análisis del contenido de los gránulos corticales ha revelado que en ellos hay una mezcla de enzimas, proteínas estructurales y materiales coloidales denominados mucopolisacáridos sulfatados. Dos de estos enzimas fueron descritos por Edward J. Carroll, cuando trabajaba en mi laboratorio (en la actualidad se encuentra en la Universidad de California en Riverside). Uno de los enzimas altera específicamente las proteínas receptoras de la membrana vitelina, de forma que los espermatozoos supernumerarios unidos a ella quedan desenganchados, y ningún otro puede tampoco adherirse posteriormente. El segundo enzima degrada las proteínas conectivas que unen la membrana vitelina con la plasmática, quedando ambas separadas entre sí. La materia coloidal liberada de los gránulos provoca por ósmosis la entrada de agua en el espacio que queda entre las membranas vitelina y plasmática, causando el hinchamiento de la membrana vitelina y su distanciamiento de la superficie del huevo.

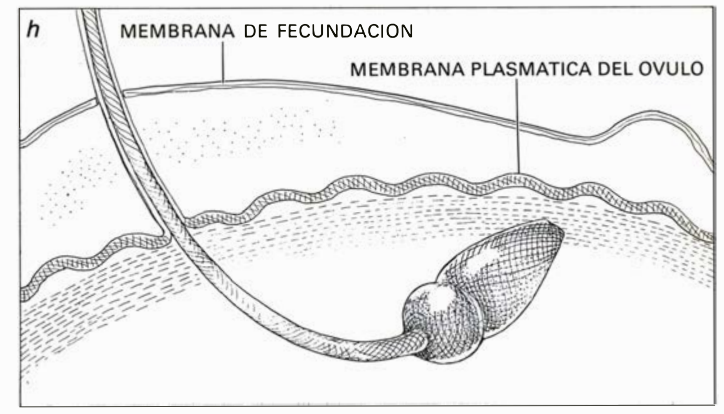
Al cabo de un minuto de contacto entre óvulo y espermatozoo, unas proteínas procedentes de los gránulos corticales se asocian con la membrana vitelina, ya desunida, transformándola en una cubierta protectora que se conoce como la membrana de fecundación. Mientras tanto, otro material coloidal procedente de los gránulos va cubriendo la superficie del huevo, y forma una membrana transparente llamada capa hialina, que más adelante habrá de representar un importante papel en la sujeción de las células formadas por la segmentación del huevo.

Toda esta secuencia de acontecimientos que siguen a la reacción cortical supone un bloqueo formidable de la polispermia. La destrucción de los lugares de anclaje de los espermatozoos y la formación de la membrana de fecundación im-









situado por debajo de la capa vitelina *(f)*. Esta reacción se propaga por toda la superficie del óvulo en 20 segundos. Un enzima liberado de los gránulos destruye los receptores de espermatozoos situados en la capa vitelina; un segundo enzima provoca la separación de la capa vitelina de la superficie del óvulo, desprendiendo los espermatozoos supernumerarios *(g)*. Las proteínas estructurales liberadas de los gránulos transforman entonces la capa separada en una cubierta protectora denominada la membrana de fecundación *(h)*. Esquema basado en otro de S. P. Schatten.

piden eficazmente el acercamiento de espermatozoos supernumerarios a la membrana plasmática y su fusión con ésta.

Si todo sucede correctamente durante la reacción cortical y después de ella, un único espermatozoo penetra en el óvulo y lo fecunda. El mecanismo por medio del cual ocurre esto último aún no se conoce, pero no depende de la motilidad del espermatozoo: el flagelo del espermatozoo deja de batir poco después de la fusión con el óvulo. Las fotografías obtenidas con microscopio electrónico de barrido por Gerald P. Schatten y Daniel Mazia, de la Universidad de California en Berkeley, sugieren que ciertas diminutas protuberancias llamadas *microvilli*, que cubren la superficie del óvulo, podrían estar implicadas en la absorción del espermatozoo anclado, quizá de manera análoga a como una ameba absorbe sus presas. Una vez que el espermatozoo ha penetrado en el óvulo, su núcleo gira 180 grados y migra hacia el núcleo de aquél. Finalmente, alrededor de 20 minutos después de la inseminación, los núcleos paterno y materno se unen. Con la condensación de los cromosonas y la primera segmentación del huevo se completa la fecundación, comenzando entonces el desarrollo embrionario.

Antes de ser fecundado, el óvulo se encuentra en un estado en el que su metabolismo está suspendido. No está muy claro el alcance exacto de tal represión, pero tanto la respiración como el transporte de substancias al exterior o al interior de la célula y la síntesis de proteínas y de ARN están considerablemente reducidos, no habiendo en absoluto síntesis de ADN. Con la fecundación se produce una activación general del aletargado metabolismo de la célula y da comienzo el desarrollo embrionario. Esta activación e iniciación no se deben a que el espermatozoo aporte algún factor del que carezca el óvulo, ya que el desarrollo de éste puede inducirse experimentalmente con sólo punzarlo mediante una aguja o bien exponiéndolo a soluciones ácidas o salinas. Aunque los embriones resultantes habitualmente no sobrevivan, ya que poseen sólo la mitad de la dotación cromosómica característica de la especie, y esto permite la expresión de los genes letales sucesivos, el hecho de que el óvulo pueda ser activado de esta manera indica que los espermatozoos se limitan a poner en funcionamiento un programa genético previamente establecido en aquél.

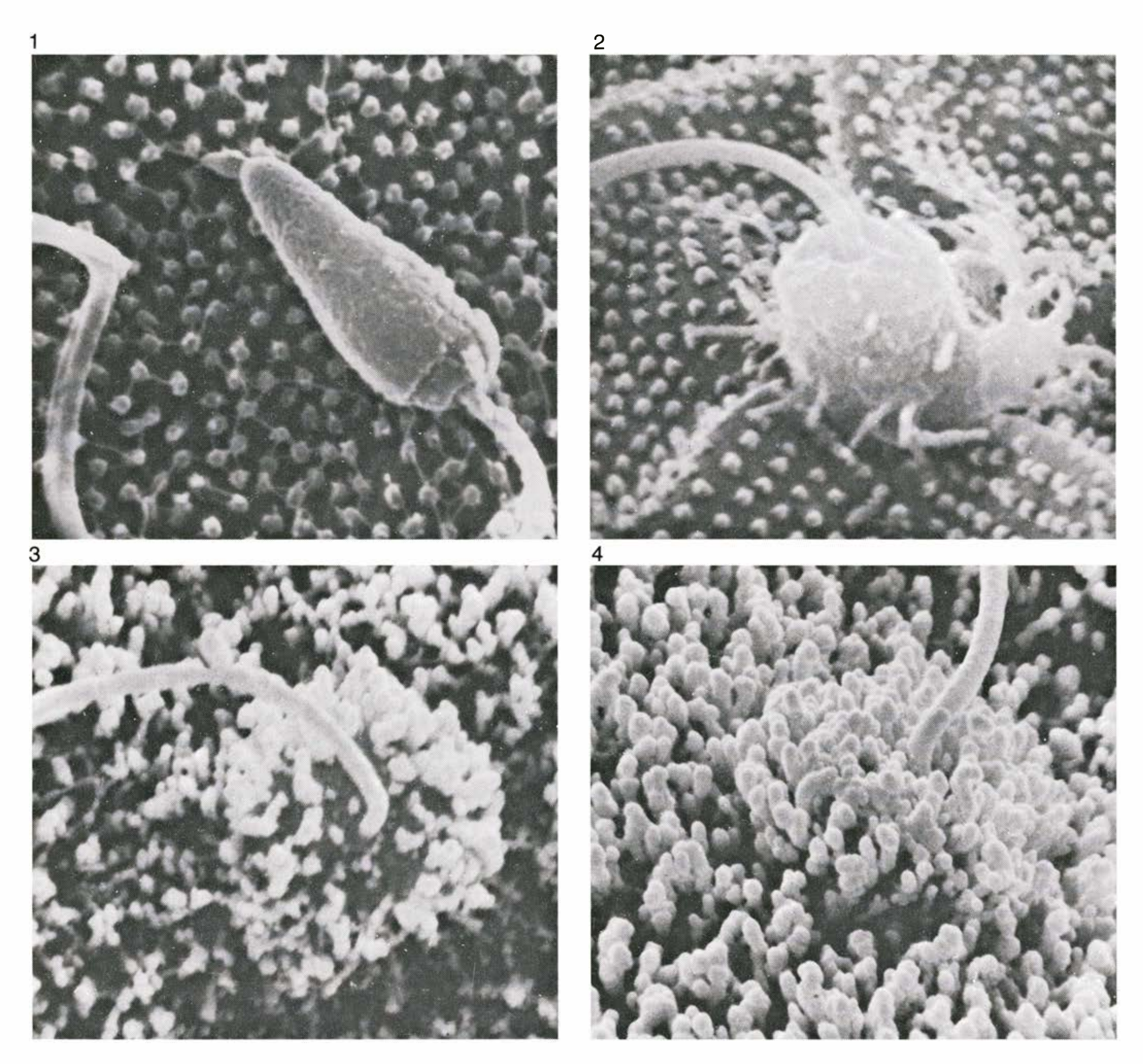
Cuando empecé a estudiar por primera vez los fenómenos bioquímicos que tienen lugar durante la activación del huevo, hace ya 13 años, quedé fascinado y a la vez confundido por el gran número de cambios que se producen en el óvulo de erizo de mar a consecuencia de la fusión con el espermatozoo. Mi investigación inicial mostró que el proceso de activación puede subdividirse en dos fases distintas: una multitud de cambios "tempranos" que ocurren durante los primeros 60 segundos a partir del contacto óvulo-espermatozoo y una serie de cambios "tardíos" que comienzan cinco minutos después del contacto.

La visión actual de los cambios que produce el proceso de la fecundación indica que los cambios tempranos empiezan con una entrada de iones sodio alrededor de tres segundos después de la adición de esperma, lo que, como hemos visto, constituye un primer bloqueo de la polispermia. Un segundo cambio, que

empieza aproximadamente a los 20 segundos, es el aumento repentino de la concentración de iones calcio en el interior de la célula. Pocos segundos después penetra el espermatozoo y da comienzo la reacción cortical. Casi simultáneamente empieza a producirse una entrada masiva en el huevo de iones sodio procedentes del agua marina y una descarga de ácido (hidrogeniones o protones) desde el huevo al agua de mar. La liberación del ácido comporta un considerable descenso de la acidez del citoplasma del huevo, como veremos más adelante.

Tres segundos después del comienzo de la reacción cortical se activa un enzima que transforma casi la mitad de la provisión celular de coenzima NAD a su forma fosforilada, NADP. Ambas substancias son formas distintas de la vitamina *B* niacina: la conversión de una en la otra desplaza el metabolismo celular a un estado más reductor, y, por ello, de mayor síntesis. Seguidamente, unos 10 a 20 segundos después del comienzo de la reacción cortical, o unos 35 a 45 después de la adición de esperma, se produce un gran aumento en el consumo de oxígeno. Casi al mismo tiempo se libera glucosa-6-fosfato deshidrogenasa desde los depósitos celulares al citoplasma. Es éste uno de los enzimas que se requieren en el metabolismo de los azúcares.

Los cambios tardíos, que empiezan cinco minutos después de la inseminación, abarcan fenómenos de biosíntesis de una gran importancia para el desarrollo embrionario. Aumenta la tasa de síntesis de proteínas, y hay un incremento en el intercambio de iones potasio a través de la membrana celular. Tam-



PENETRACION DEL ESPERMATOZOO FECUNDANTE en el óvulo. Primero, el espermatozoo se fija a la capa vitelina por medio del filamento acrosómico *(1)*; seguidamente, se fusiona con la membrana plasmática. El óvulo responde con una prolongación localizada de los microvilli, que empiezan a rodear al espermatozoo *(2)*. Luego tiene lugar la reacción cortical, que provoca la separación de la capa vitelina y su transformación en la membrana de fecundación. Las dos últimas imágenes se obtuvieron retirando del óvulo la membrana de fecundación. Tres minutos después de la fecundación, el espermatozoo se ha medio incorporado al óvulo y está recubierto por su membrana plasmática *(3)*, un minuto más tarde, sólo se aprecia la cola del espermatozoo *(4)*. Las microfotografías *1*, *3* y *4* han sido obtenidas por Tegner; la *2* por Schatten y Daniel Mazia, de la Universidad de California en Berkeley. Las fotografías muestran un aumento de 15.000 diámetros.

bién a los cinco minutos se activan los sistemas de transporte que introducen aminoácidos, fosfato y nucleósidos en la célula: se forman así los distintos bloques que componen la estructura de las proteínas y del ADN. A los 20 o 25 minutos empieza el primer ciclo de síntesis de ADN.

¿Cómo puede ser suficiente la fusión de un espermatozoo con sólo el 0,0002 por ciento de la superficie del óvulo para desencadenar todos estos cambios? Una explicación puede ser que la diminuta perturbación causada por el espermatozoo queda amplificada en la superficie del óvulo, o cerca de ella, y que la señal amplificada se transmite entonces a la maquinaria bioquímica del citoplasma. En otros sistemas, los mensajeros intracelulares como el AMP cíclico y los pequeños iones como el calcio funcionan de la misma manera. Los fenómenos de activación, sin embargo, no han sido correlacionados con un aumento de los niveles de AMP cíclico.

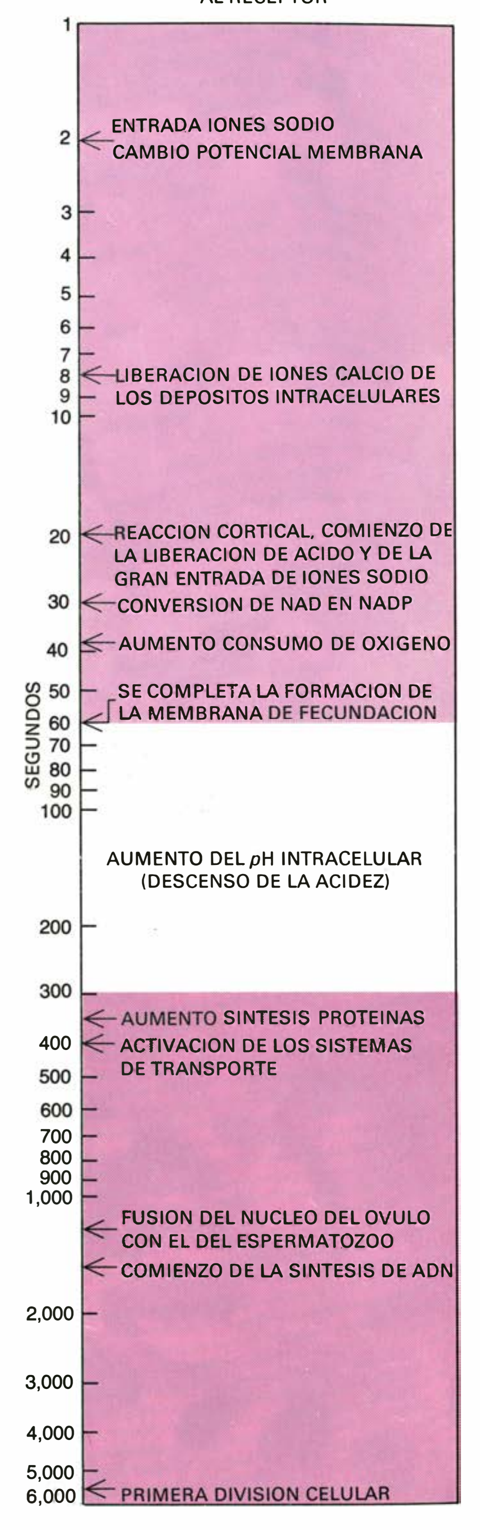
En 1937, mi antiguo profesor Mazia, que por aquel entonces trabajaba en el laboratorio de L.V. Heilbrunn de la Universidad de Pennsylvania, observó por primera vez que la concentración de iones calcio del interior de la célula aumenta después de la fecundación. Hoy en día los hallazgos de Mazia se considerarían sospechosos, ya que el incremento de los niveles de calcio se había detectado en homogeneizados de huevos preparados en distintos tiempos después de la fecundación. Quizá, podría argumentarse, el incremento observado es simplemente un artefacto causado por la destrucción de las células.

La validez de las observaciones de Mazia ha sido verificada, de todas formas, en un experimento llevado a cabo recientemente por Ellis B. Ridgway, del Medical College of Virginia, en colaboración con John C. Gilkey y Lionel Jaffe, de la Universidad de Purdue. Emplearon la proteína luminiscente acuarina, que se extrae de una medusa. La acuarina brilla sólo en presencia de iones calcio, de forma que es de gran utilidad para detectar los cambios en la cantidad de calcio libre del interior de las células. Ridgway y sus colaboradores observaron un bajo nivel de luminiscencia al inyectar acuarina en los óvulos de gran tamaño del pez *Oryzias latipes*. Un minuto después de la fecundación, la luminosidad aumentaba diez mil veces, lo que indicaba una liberación de grandes cantidades de calcio en el interior de la célula. La luminiscencia permanecía elevada durante el período de la reacción cortical, descendiendo después al nivel original. Si los investigadores permanecían previamente en una habitación a obscuras para adaptar sus ojos, y provocaban entonces la fecundación de un único óvulo, podían apreciar su brillo incluso a simple vista.

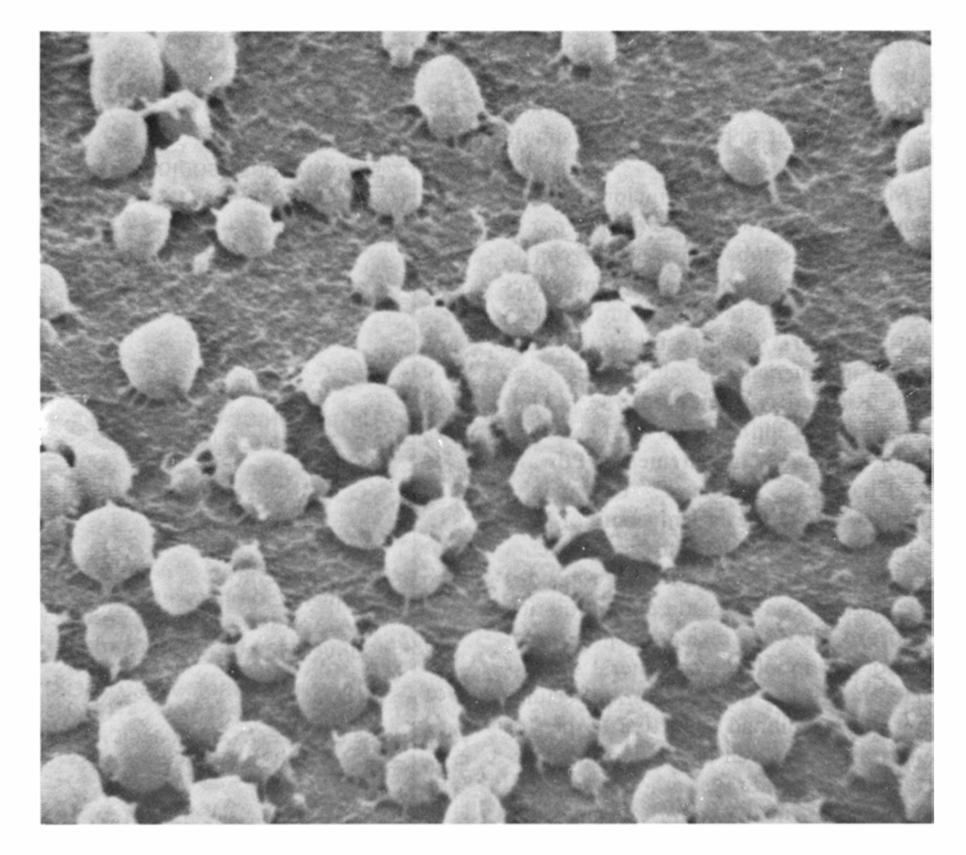
Mazia sugirió en principio que el aumento de calcio libre inducido por la fecundación estaba relacionado con la activación del huevo, pero sin saber exactamente si este incremento era una causa primaria de la activación o sólo una consecuencia de ella. Hace unos años Richard A. Steinhardt, de la Universidad de California en Berkeley, y el autor, e independientemente Edward L. Chambers, Berton C. Pressman y Birgit L. Rose, de la University of Miami School of Medicine, se dieron cuenta de que podían distinguir entre estas dos alternativas con la ayuda del antibiótico ionóforo A23187, recientemente descubierto. Los ionóforos en general son fármacos que dan a la membrana celular una permeabilidad selectiva a ciertos iones. El A23187 en particular hace permeable la membrana a iones con doble carga positiva, como el calcio o el magnesio. Si el incremento de calcio por sí solo causaba la activación del huevo, razonamos entonces que la presencia de este ionóforo debería desencadenar el desarrollo partenogenéticamente (en ausencia de espermatozoos).

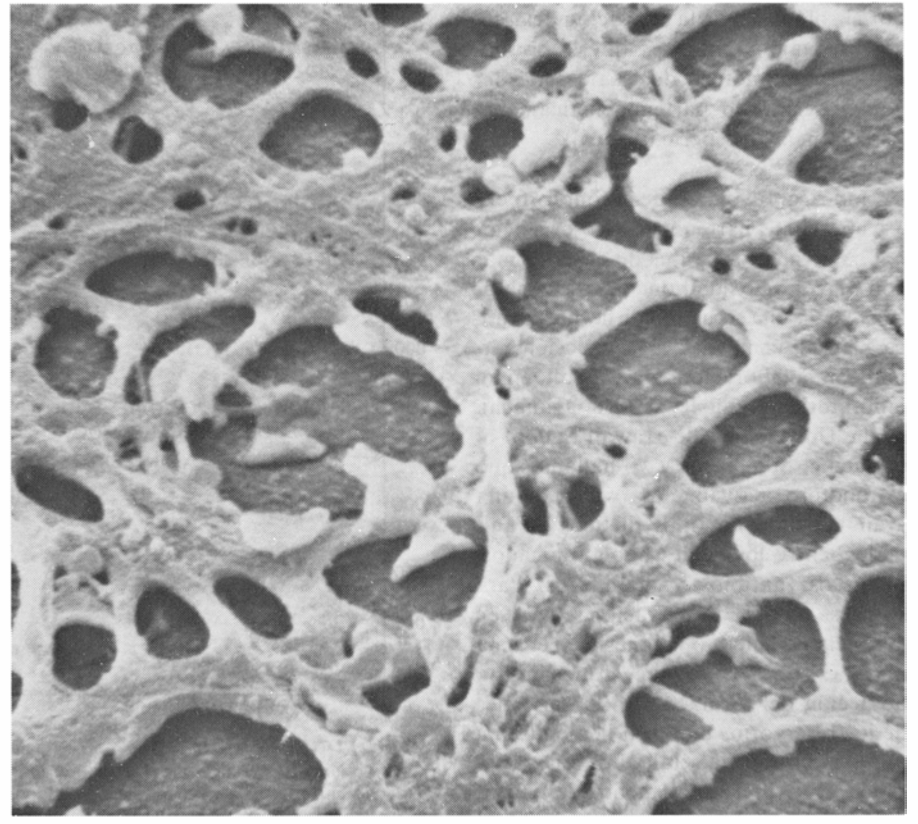
Ambos obtuvimos que el ionóforo era un agente partenogenético excepcional. Activaba los óvulos de erizo de mar con mayor velocidad que los espermatozoos y con una secuencia de hechos idéntica a la que acompaña a la fecundación normal. En colaboración con Carroll, en mi laboratorio, y con Ryuzo Yanagimachi, de la Universidad de Hawaii School of Medicine, Steinhardt y yo hallamos que el ionóforo activaba los óvulos de una gran variedad de tipos filéticos de animales, entre los cuales había tunicados, moluscos, anfibios y mamíferos. Para sorpresa nuestra, la activación se daba en ausencia de calcio o magnesio en el agua marina; ello sugería que el mecanismo de acción del fármaco era liberar iones de los reservorios unidos a la membrana por el interior de la célula. Los iones liberados eran presumiblemente de calcio, ya que la mayoría de iones magnesio del óvulo del erizo de mar se encuentran en estado libre, mientras que los iones calcio se encuentran en forma combinada de la que pueden ser liberados.

¿Qué es exactamente lo que hace el calcio? Un elegante experimento, llevado a cabo por Vacquier, indica que el calcio induce directamente la fusión de los grá-



CRONOLOGIA de los fenómenos que siguen a la fecundación del óvulo del erizo de mar *Strongylocentrotus purpuratus*, representada en escala semilogarítmica. Pueden dividirse en una fase temprana, que tiene lugar en los primeros 60 segundos después de la fecundación *(color claro)*, y en una fase tardía, que empieza alrededor de 300 segundos después de la fecundación *(color obscuro)*. En los cambios tempranos están principalmente implicadas moléculas pequeñas, como iones y coenzimas; los cambios tardíos comprenden la síntesis de macromoléculas, como proteínas y ácido desoxirrobunucleico, así como la activación de los sistemas de transporte de aminoácidos y nucleósidos.



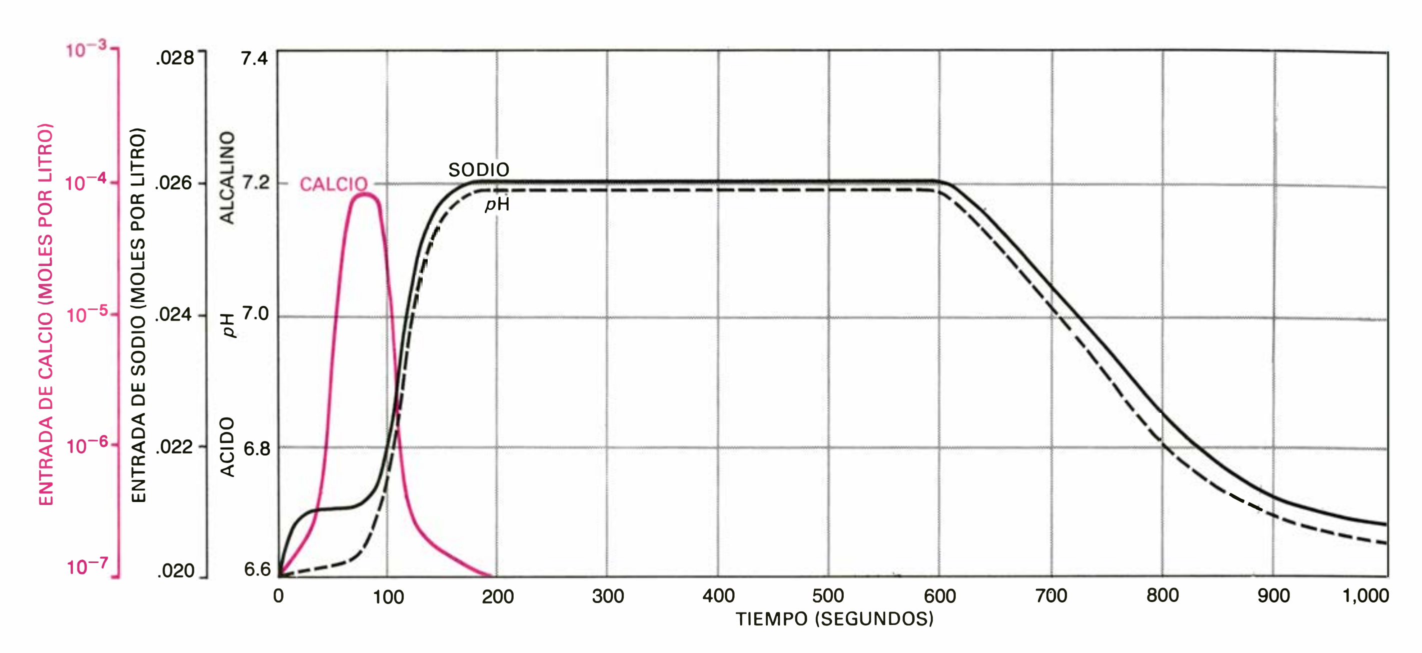


LOS GRANULOS CORTICALES AISLADOS descargan su contenido mediante la adición de iones calcio. Victor D. Vacquier, de la Universidad de California en Davis, aisló los gránulos corticales fijando óvulos de erizo de mar en un portaobjetos y rompiéndolos con un chorro de agua sin calcio. El chorro se llevaba el citoplasma, quedando la membrana plasmática "vuelta del revés" con unos cuantos gránulos corticales adheridos a ella, como se muestra en la fotografía electrónica de barrido en la parte superior. Al añadir iones calcio a esta preparación, los gránulos se unían unos a otros en una reacción explosiva *(abajo)*. Este experimento sugiere que el proceso normal de fusión de los gránulos corticales con la membrana plasmática resulta de un aumento de los iones calcio que se encuentran libres en el interior del óvulo. El aumento es de 10.000 diámetros.

nulos corticales con la membrana plasmática del óvulo. Vacquier fijó óvulos de erizo de mar a un portaobjetos de microscopio que había sido recubierto con una proteína "pegajosa" cargada positivamente, como la protamina. Sometió entonces las células fijadas a un chorro de agua sin calcio, provocando su rotura, y quedando retenida su membrana plasmática junto con los gránulos corticales asociados a ella. Vacquier observó que los gránulos corticales de estas células "vueltas del revés" liberaban su contenido cuando se les añadía calcio.

¿Provoca la reacción cortical, inducida por calcio, el desencadenamiento de los cambios tardíos, como la síntesis de proteínas y de ADN? Diversas observaciones importantes indican que no es así. Por ejemplo, muchos fármacos que bloquean la reacción cortical no impiden la activación del desarrollo embrionario. Es más, en 1972, Steinhardt y Mazia descubrieron que incubando óvulos en bajas concentraciones de amoníaco podían activar varios de los cambios tardíos, como la síntesis de proteínas, de ADN e incluso la condensación de los cromosomas, sin haber activado los cambios tempranos, como la reacción cortical. Experimentos posteriores efectuados por Carroll, en mi laboratorio, y por Vacquier mostraron que la incubación de óvulos en una variedad de compuestos que tenían un grupo amino (NH_2), por ejemplo la procaína anestésica, también activaban los cambios tardíos sin haberse presentado los tempranos. Todos estos hallazgos sugieren que los cambios tempranos no son una condición previa necesaria para los tardíos.

El primer indicio sobre el mecanismo de activación de los cambios tardíos los obtuvimos en nuestra investigación de la descarga de ácido por parte del óvulo, que comienza al principio de la reacción cortical y persiste durante los cuatro minutos siguientes de desarrollo. Un ácido se define como la molécula que cede fácilmente un protón a otra molécula: puede ser también un protón libre (un átomo de hidrógeno sin su único electrón). Es decir, el óvulo va soltando moléculas ácidas, o bien protones, al agua marina que le rodea. En principio pensamos que se liberaban moléculas ácidas, ya que se sabía que los mucopolisacáridos sulfatados, descargados de los gránulos corticales, eran ácidos. No obstante, surgió un inconveniente con esta hipótesis cuando Miles R. Paul, que entonces trabajaba en mi laboratorio como estudiante graduado, descubrió



VARIACIONES EN LA CONCENTRACION de iones sodio, calcio e hidrogeniones (protones) que muestran hallarse implicados en muchos de los cambios inducidos por la fecundación. Pocos segundos después de la fecundación hay una pequeña entrada de iones sodio en el huevo, asociada con una despolarización eléctrica transitoria de la membrana celular. A los 20 segundos se liberan iones calcio de depósitos intracelulares, que aparentemente desencadenan la reacción cortical y otros cambios tempranos. Alrededor de los 60 segundos, el contenido en sodio del huevo empieza a incrementarse, junto con una salida de protones y un aumento del *p*H, que de alguna forma induce la actividad y la división celular. A los 180 segundos la concentración de calcio ha descendido al nivel original, y, entre los 10 y los 20 minutos, el sodio y el *p*H retornan a sus niveles primitivos.

que los óvulos del gusano equiuroideo *Urechis caupo,* que no sufren reacción cortical, también liberaban ácido al ser fecundados. Surgió un segundo problema con el ácido que se originaba en los gránulos corticales, al calcular la cantidad de ácido desprendido por los óvulos del erizo de mar: si el ácido estaba contenido sólo en los gránulos, habría de estar a una concentración de 0,1 moles de ácido por litro, que parece ser demasiado elevada para un funcionamiento normal de la célula.

La prueba definitiva de que la liberación de ácido no estaba relacionada con la reacción cortical provino de nuestro análisis de la activación de los cambios tardíos por medio de amoníaco y otras aminas. James D. Johnson y el autor, en la Scripps Institution, y Paul, que trabajaba en la Universidad de Victoria, en el Canadá, hallaron que se liberaba ácido cuando, al poner los óvulos en una solución de amoníaco o procaína, se desencadenaban los cambios tardíos, incluso sin haber tenido lugar la reacción cortical. Cuando añadimos después espermatozoos para inducir la reacción cortical, observamos con sorpresa que no había liberación adicional de ácido. Por tanto, la liberación de ácido está relacionada con la activación de la síntesis y no con la secreción de los gránulos corticales.

La siguiente pista sobre la naturaleza de la activación fue el descubrimiento de Chambers de que se requieren iones sodio para proseguir el desarrollo, después de la reacción cortical. Al fecundar óvulos en agua de mar normal y transferirlos luego a agua de mar sin sodio, los núcleos del espermatozoo y del óvulo no lograron unirse, y el desarrollo se detuvo. Daba la impresión de que los óvulos habían detenido su actividad. Cuando Chambers añadía iones sodio a la solución, el desarrollo proseguía allí donde se había detenido. El requerimiento de sodio se mantenía sólo durante unos pocos minutos después de la reacción cortical. Si los óvulos eran fecundados en agua de mar normal y transferidos a agua de mar sin sodio diez minutos después, el desarrollo no se detenía.

Este descubrimiento sugirió que el sodio estaba implicado en la activación biosintética. ¿Estaba relacionado con la liberación de ácido? Para hallar la respuesta detuvimos primero el desarrollo de los óvulos, poniéndolos en agua de mar sin sodio, 60 segundos después de la fecundación, y medimos entonces la acidez de la suspensión. Como ya esperabamos, no se había liberado nada de ácido. No obstante, al añadir iones sodio a la suspensión aparecía ácido en el agua marina y los huevos empezaban a desarrollarse. Hicimos el experimento con diversos iones que normalmente se hallan en el agua de mar, pero sólo el sodio iniciaba la liberación de ácido y estimulaba el desarrollo.

Esta observación sugirió que los iones sodio estaban implicados en un intercambio sodio-protón, en el cual un ion de sodio entra en la célula por cada protón que sale de ella. Esta hipótesis se apoyaba en nuestro hallazgo de que, al medir la absorción de sodio con el isótopo radiactivo sodio 22, ésta se correspondía temporal y cuantitativamente con la liberación de ácido. También encontramos que al bloquearse la entrada de sodio por medio del fármaco diurético amiloride, no había activación del desarrollo. Por tanto, la absorción de sodio y la liberación de ácido parecía ser esencial para el comienzo de la embriogénesis.

A continuación nos planteamos cuál de estos dos procesos era el esencial para el desencadenamiento del desarrollo. Un experimento sugerido por los trabajos de Chambers nos dio la respuesta. El había encontrado que si se detenía el desarrollo 60 segundos después de la fecundación poniendo el huevo en agua marina sin sodio, podía luego ponerse en marcha de nuevo simplemente añadiendo amoníaco al agua salina. Descubrimos nosotros que en estas condiciones había liberación de ácido, lo que indica que es la salida de protones y no la entrada de iones el hecho crucial que pone en marcha el desarrollo.

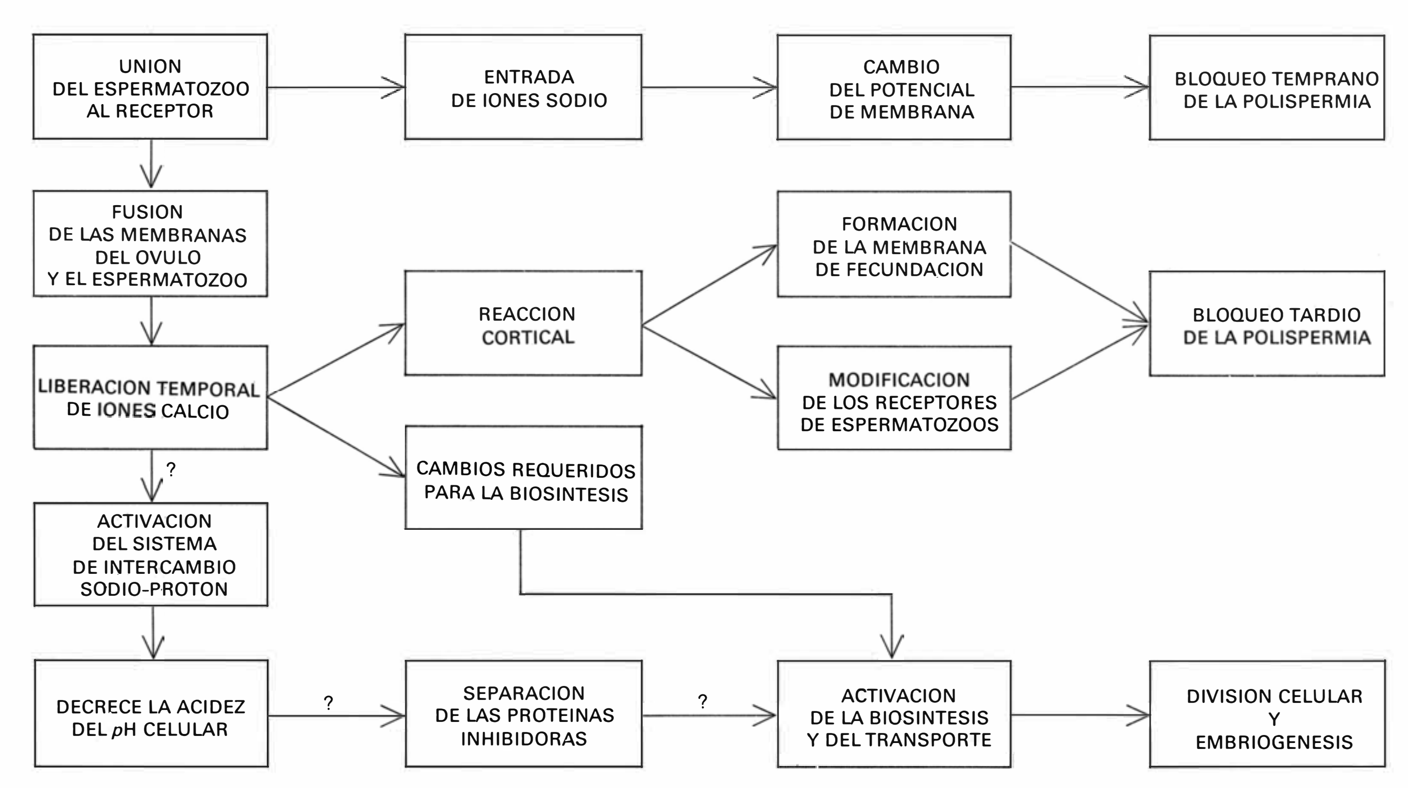
¿Cuál es el efecto de la salida de ácido en el huevo? La hipótesis más razonable es que la pérdida de ácido aumenta la alcalinidad del citoplasma del huevo. Comprobamos esta hipótesis con el mé-

todo poco ortodoxo, pero efectivo, de homogeneizar los huevos a tiempos distintos tras la fecundación, y medir directamente la acidez mediante un electrodo de *p*H. Suponiendo que estas mediciones pueden extrapolarse directamente a la célula viva, revelan que el *p*H aumenta (se hace menos ácido) en 0,4 unidades de *p*H durante los primeros cinco minutos después de la fecundación. Este desplazamiento del *p*H ha sido confirmado recientemente por Sheldon Shen en el laboratorio de Steinhardt, que ha trabajado con microelectrodos insertados directamente en el huevo. Todas las pruebas disponibles indican que es la variación de *p*H lo que pone en marcha la maquinaria de síntesis del huevo después de la fecundación.

Este descubrimiento ha dado pie a dos cuestiones fundamentales. Una, ¿cuál es el cambio temprano dependiente del calcio que activa el sistema de intercambio sodio-protón? La resolución de esta pregunta requiere el conocimiento de cómo modifican los cambios tempranos la estructura de la superficie del huevo. Una posibilidad es que la reacción cortical inducida por el calcio altere la superficie del huevo de tal forma que el sistema de intercambio sodio-protón quede expuesto o activado. Una segunda posibilidad es que el aumento de la concentración de calcio en el interior de la célula altere directamente la permeabilidad de la membrana plasmática del huevo. Esta hipótesis se desprende de los experimentos de Michael Berridge y Robert Meech, de la Universidad de Cambridge, que hallaron que una elevada concentración de iones calcio en el interior de la célula aumenta la permeabilidad de las células de las glándulas salivares de insectos y de las células nerviosas de caracoles. Una tercera posibilidad es que la liberación de calcio afecte a proteínas contráctiles, como la miosina, de tal forma que el sistema de intercambio sodio-protón se disponga en la configuración adecuada para ser activo. Los futuros experimentos deberán permitirnos determinar cuál de estas tres posibilidades es la correcta.

La otra cuestión es: ¿Cómo logra el aumento de *p*H poner en marcha procesos metabólicos tales como la síntesis de proteínas y la replicación del ADN? Los recientes trabajos de David Nishioka, un investigador postdoctorado de mi laboratorio, y de Shen y Steinhardt han mostrado que la variación de *p*H inducida por la fecundación es transitoria, y rápidamente desciende al nivel registrado en el óvulo. Nishioka ha hecho el importante descubrimiento de que, para iniciar el desarrollo, el *p*H sólo necesita aumentar durante un período de diez minutos. Este hallazgo significa que el cambio de *p*H no actúa directamente incrementando la tasa de procesos sintéticos, ya que la mayoría de éstos tienen lugar después del aumento temporal de *p*H. Antes, esta variación transitoria de *p*H debe iniciar un cambio en la célula que, a su vez, tiene como consecuencia la activación del huevo.

¿Cuál puede ser este cambio sensible al *p*H? Dado que son muchas las transformaciones que comienzan durante la activación, puede haber numerosos procesos sensibles al *p*H, o bien, alternativamente, un efecto generalizado y persuasivo del *p*H aumentado en la célula. Una hipótesis atractiva, que apoyaría este efecto general del *p*H en la célula, es que una molécula inhibidora se una a proteínas muy determinadas, como enzimas o proteínas ribosómicas, y que la variación temporal de *p*H disociara al inhibidor de estas proteínas, permitiendo su funcionamiento en reacciones sintéticas.



MAPA DEL FLUJO que pone de manifiesto las hipotéticas relaciones entre los cambios tempranos dependientes del calcio (como la reacción cortical) y los cambios tardíos dependientes del *p*H (como la síntesis del ADN). El mecanismo por el cual el calcio pone en marcha el sistema de intercambio sodio-protón que provoca una disminución de la acidez del citoplasma del huevo (un aumento del *p*H) no se conoce todavía. Tampoco se sabe cómo esta variación del *p*H "pone en marcha" la síntesis de proteínas y de ADN, pero podría estar implicada en ello la supresión de proteínas inhibidoras de algunas proteínas funcionales, como las de los ribosomas, o de proteínas estructurales, como las subunidades de actina o tubulina. En este último caso, la supresión de una proteína inhibidora permitiría polimerizarse a las subunidades, formando filamentos, lo que tendría como consecuencia importantes cambios en la estructura celular; ello podría dar comienzo a una gran variedad de procesos bioquímicos.

Una vez que los inhibidores hubieran sido destruidos, el huevo estaría irreversiblemente determinado a desarrollarse.

Un ejemplo interesante de este tipo de regulación, y que bien pudiera aplicarse a la activación del huevo, ha sido descrito por Lewis G. Tilney, de la Universidad de Pennsylvania. Tilney ha estudiado la formación del filamento acrosómico durante la fase inicial de la fecundación, y ha descubierto que el filamento se crea por la polimerización explosiva de actina, una proteína contráctil que puede hallarse en dos estados distintos: como subunidades individuales o como cadena filamentosa. Hasta que en el espermatozoo se desencadene la formación del filamento acrosómico, la actina de la cabeza permanece despolimerizada, aparentemente por medio de la asociación de las subunidades de actina con una proteína inhibidora. Cuando el espermatozoo se activa con el contacto con la cápsula gelatinosa del óvulo, la proteína inhibidora se disocia y las subunidades de actina se polimerizan espontáneamente. ¡Tilney y sus colegas han demostrado recientemente que la disociación de la proteína inhibidora se debe a un incremento del *p*H intracelular! Aplicándolo a la activación del huevo, puede suponerse que ciertas proteínas inhibidoras están unidas de forma análoga a las principales proteínas estructurales del huevo, como la actina o quizá la tubulina. La variación de *p*H provocaría, por tanto, la disociación de las proteínas inhibidoras unidas a las subunidades de actina o tubulina, permitiendo su polimerización. Este hecho traería como consecuencia una amplia reorganización de la estructura celular y daría comienzo a una variedad de cambios, aparentemente no relacionados.

Una cuestión fundamental que se deriva de estos descubrimientos es si los mecanismos iónicos que regulan la actividad celular en los comienzos de la vida siguen haciéndolo en los estadios posteriores de la existencia del organismo. El tejido nervioso y el muscular son ejemplos conocidos en los que la función del tejido está estrechamente relacionada con pequeños iones, como los de sodio, potasio y calcio, y existen indicios recientes de que los iones podrían participar también en la regulación de procesos de mayor complejidad, como la división y la diferenciación celular. Un gran número de investigadores han comprobado que incrementando los niveles intracelulares de iones calcio o magnesio, bien por medio de ionóforos, bien simplemente añadiendo los iones al medio, logran transformar las células de cultivos de tejidos, en principio en reposo, en células con una gran actividad divisoria.

Más sorprendente aún es la reciente demostración de Clarence D. y Charlotte Cone, del Veterans Administration Hospital de Hampton, Virginia, de que puede inducirse la división de células nerviosas (que normalmente no se dividen nunca) incrementando la concentración de iones sodio en su interior.

Los cambios que se producen durante el desarrollo también parecen estar relacionados con variaciones iónicas. Por ejemplo, Lester G. y Lucena J. Barth, trabajando en Woods Hole, Massachusetts, han hallado que pequeñas variaciones en la concentración de iones sodio pueden desviar la diferenciación de las células epiteliales de embriones de anfibios de forma que lleguen a dar células nerviosas. Un segundo ejemplo es la inducción de procesos de desarrollo en vegetales por medio de la luz. El efector es el pigmento vegetal fitocromo, y los trabajos de Ian A. Newman y Winslow R. Briggs, de la Carnegie Institution of Washington's Department of Plant Biology, ha mostrado que el primer cambio inducido por la luz afecta a la permeabilidad iónica de la membrana celular.

No está muy claro cómo los iones de pequeño tamaño pueden producir efectos tan diversos, y, desde luego, la dificultad que supone imaginar posibles mecanismos ha conducido a muchos investigadores a pasar por alto o a despreciar las hipótesis de que la modulación del contenido iónico pudiera ser un factor importante en la regulación de las funciones celulares, o un iniciador de cambios durante la diferenciación celular. El descubrimiento de que cambios esporádicos en el calcio intracelular y en el *p*H intracelular desencadenan el desarrollo de los huevos del erizo de mar ofrece un ejemplo que pone de manifiesto cómo los iones pueden poner en marcha actividades celulares completamente nuevas. No sabemos aún si el *p*H intracelular es un regulador importante del desarrollo en huevos de otras especies, o si está implicado en otras funciones celulares. No obstante, parece probable que la regulación por medio de modulaciones esporádicas o permanentes del contenido iónico esté extendida por todo el mundo animado y que los mecanismos que regulan la actividad celular en la fecundación se emplean a lo largo de toda la vida del organismo.

Consecuencias de una anterior crisis energética

En el siglo XVI, Inglaterra se quedó sin madera y recurrió al carbón. La adopción del nuevo combustible originó una cadena de hechos que culminaron, dos centurias más tarde, en la Revolución Industrial

John U. Nef

En la Europa medieval la madera se usaba no sólo en muchos tipos de construcción sino también como fuente principal de calor en los hogares y en la industria. Luego, en Inglaterra, en la segunda mitad del siglo XVI, la hulla llegó a emplearse ampliamente como combustible sucedáneo de la madera. La primera economía energizada por carbón que el mundo ha conocido se estableció primero en Inglaterra y luego en Escocia entre 1550 y 1700, aproximadamente. Esta transición desde la tala de árboles hasta la minería del carbón como fuente térmica principal formó parte de una temprana revolución económica inglesa. La primera crisis energética, que tiene muchos aspectos en común con la que hoy nos enfrentamos, fue una crisis de desforestación. La adopción del carbón cambió la historia económica de Inglaterra, luego la del resto de Europa y, finalmente, la del mundo. Condujo a la Revolución Industrial, que se puso en marcha en Inglaterra en las dos últimas décadas del siglo XVIII. La sustitución de la madera por el carbón entre 1550 y 1700 llevó a nuevos métodos de fabricación, a la ampliación de las industrias existentes y a la explotación de recursos naturales no abordados hasta entonces.

Hacer estas afirmaciones no equivale a menospreciar la función desempeñada por otros cambios, ocurridos en la Edad Media y el Renacimiento, en el nacimiento de nuestro mundo industrializado. El siglo anterior a la crisis maderera inglesa –los años comprendidos entre 1450 y 1550– se caracterizó por un nuevo espíritu de expansión. Se lanzaron viajes de descubrimiento que llevaban exploradores a los confines de la tierra. El arte de la impresión con tipos movibles se difundió por toda Europa, y aumentó la producción de papel; se imprimieron y pusieron en circulación millones de libros. En Europa Central, en donde se hallaban los principales centros mineros y metalúrgicos, el rendimiento de minerales, especialmente los minerales de cobre argentíferos, se multiplicó por un factor de varias unidades. A los años comprendidos entre 1494 y 1529 se les ha descrito como una época en que advino "una revolución en el arte de la guerra". Con ayuda de las nuevas armas de fuego, España conquistó México y Perú.

Estas y otras innovaciones incrementaron, directa o indirectamente, la necesidad de todos los tipos de energía existentes: el calor proporcionado por la madera y la fuerza proporcionada por el viento, los animales y las corrientes de agua. La necesidad de mayores cantidades de madera para construcción y calefacción, en especial para la fundición y refino de minerales, exigía un aumento considerable de la tala de árboles.

En toda Europa se hicieron sentir estas presiones, y no obstante fue Inglaterra la primera zona de cierta extensión en la que se experimentó una aguda escasez de madera. ¿Por qué la revolución de carburantes que condujo a nuevos usos de la energía térmica empezó precisamente allí? ¿Era especialmente escasa la madera en ese lugar? Parece ser cierto que las partes más populosas de Escocia (las zonas que rodean el estuario del Forth) carecían de forestación; se dice que un ingenioso inglés había comentado, durante el reinado de Jacobo I, que si Judas se hubiese arrepentido en la patria del rey (Escocia), ¡se habría visto en aprietos para hallar un árbol en donde colgarse! Esa explicación no calza con Inglaterra, en donde la crisis maderera ha de atribuirse a las necesidades de una agricultura, una industria y un comercio en expansión, estimulados por una población creciente y cambiante.

Parece que Suecia y Holanda fueron los únicos otros países europeos que experimentaron algo comparable al crecimiento de la población inglesa desde 1550 a 1700. La población de Inglaterra y Gales, de unos tres millones al empezar la década de 1530-39, se había casi duplicado hacia 1690. La consecuente demanda de madera para diversos fines aumentó aún más debido a cambios en la distribución demográfica. En dicho período, la población de Londres se multiplicó a lo menos por ocho, de unas 60.000 personas en 1534 a unas 530.000 en 1696.

Según el cálculo hecho por Gregory King para este último año, la capital británica había llegado a ser la mayor ciudad de Europa y quizá del mundo. King calcula que las demás "ciudades y... plazas" de Inglaterra tenían una población total de unas 870.000 personas. Esto significa que si en la década de 1530-39 sólo uno de cada diez era "conciudadano", en la década de 1690 lo era uno de cada cuatro. Unas ciudades de mayor tamaño suponían una demanda más intensa sobre los abastecimientos de madera cercana. Además, fuera de las ciudades, existía una importante migración a través del país, de desempleados en busca de trabajo. Dondequiera que hallasen empleo, habían de contar con un albergue, lo que significaba una carga más para los bosques.

Durante los reinados de Isabel I (1558 a 1603) y Jacobo I (1603-25), esta presión sobre el abastecimiento de árboles se reflejó en el coste, siempre mayor, de la leña y de la madera para construcción. El período comprendido entre 1550 y 1640 fue una época de inflación en toda Europa, pero el precio de la madera en Inglaterra aumentó con una rapidez mucho mayor que el de cualquier otra mercancía de uso general en cualquier parte. De todos los rincones del reino llegaban quejas sobre la desforestación.

En todos los lugares de Europa en

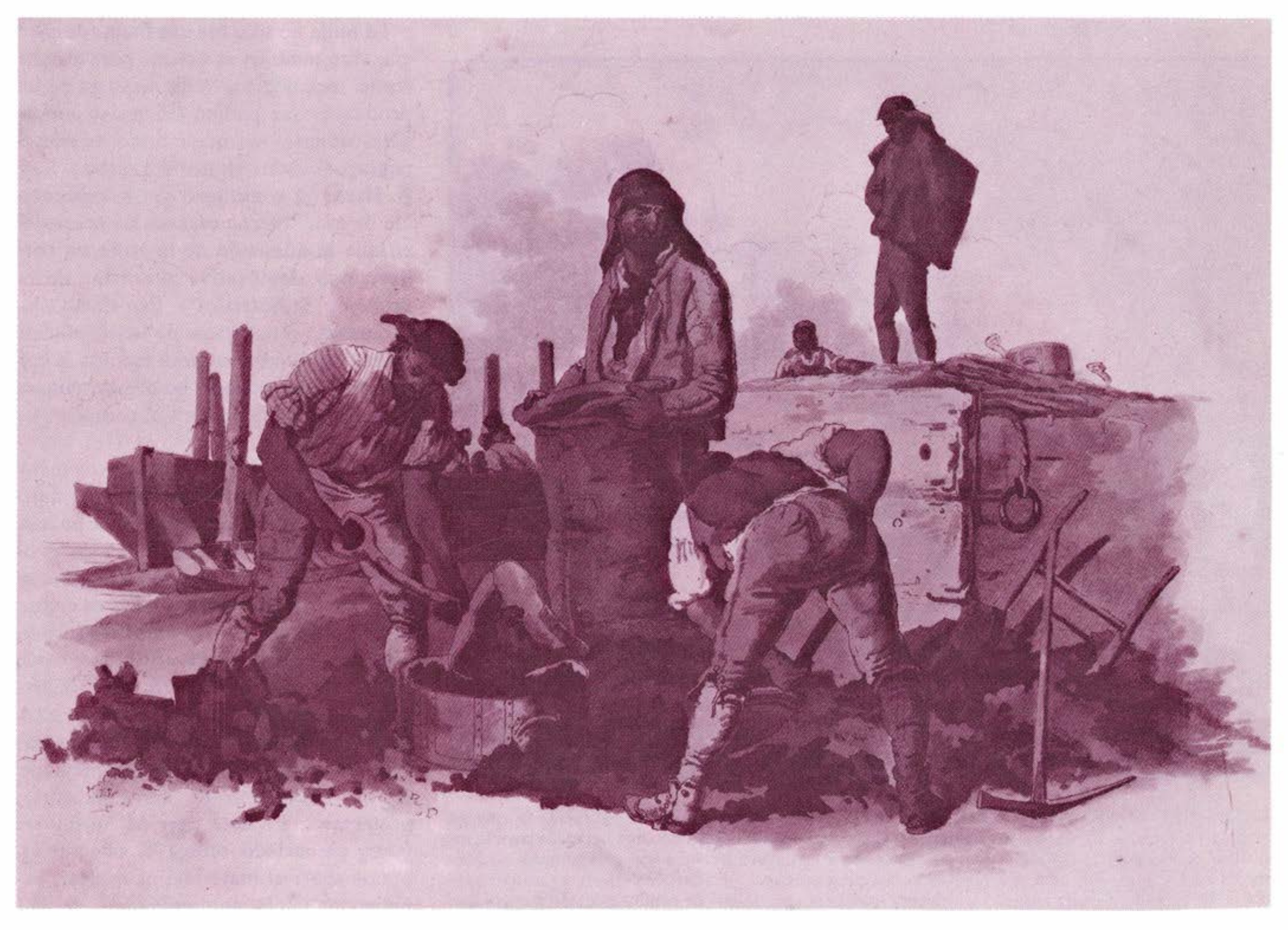
donde afloraban capas de hulla, se había quemado ésta en pequeñas cantidades desde el siglo XII. (Su uso había sido más difundido ya antes de entonces en China y también, en cierta medida, en la Britannia romana.) Durante las postrimerías de la Edad Media los campesinos europeos habían calentado ocasionalmente sus hogares o habían cargado sus hornos y fraguas con estas "piedras negras". ¿Por qué, entonces, no se generalizó el carbón como combustible en el continente y en Inglaterra, antes de que se produjese un grave agotamiento de los bosques?

En sociedades anteriores a la que surgió en Europa occidental en la época medieval, se miraba con malos ojos a la minería. Se la consideraba como un robo, incluso como una especie de violación. Al revés del arado, que hacía fértil la tierra, el pico y la pala sacaban lo que se suponía ser un suelo y subsuelo irremplazables.

Al iniciarse el siglo XVI, una actitud diferente hacia la explotación de los recursos subterráneos más valiosos halló expresión en dos libros. En *De re metallica* (1556), Georgius Agricola (1494-1555) consideraba la vocación del minero más elevada que "la del mercader que comercia por afán de lucro". Y en *Pirotechnia* (1540), Vannoccio Biringuccio (1480 a 1539) propugnaba un asalto global a estas riquezas subterráneas. Aconsejaba "a todo aquel que explote minas... perforar en el centro de las montañas... como si se tratase de obra de necromancia o de gigantes. No sólo deberían quebrantar y abrir las montañas, sino volver del revés su médula misma, a fin de que pueda verse lo que está dentro y expoliar lo antes posible la dulzura del fruto".

La nueva dignidad adscrita a la minería se reservaba para los minerales metálicos. No se extendía al carbón. Los artesanos medievales que necesitaban combustible querían que su obra fuese hermosa, estuviese destinada ya sea a sus iglesias o a seglares ricos. Los desagradables humos y vapores de carbón limitaban por tanto su mercado. Antes de mediados del siglo XVI había escasos incentivos para excavar profundamente el suelo en búsqueda de este combustible sucio mientras hubiese madera disponible, y parecía haberla en abundancia. El propio Biringuccio creía que los bosques europeos podían satisfacer todas las concebibles demandas futuras de combustible. Escribió así en *Pirotechnia:* "Es más probable que los mineros agoten los minerales disponibles, que los leñadores la madera necesaria para fundirlos. En todas partes hay grandes bosques, lo que nos hace pensar que las edades del hombre jamás los consumirían... especialmente por cuanto Natura, tan pródiga, produce otros nuevos cada día". En su voluminoso tratado se menciona la hulla sólo una vez y, aún así, para descartarla: "Además de los árboles, las piedras negras, que aparecen en muchas partes, tienen la naturaleza del carbón, pero la abundancia de árboles hace innecesario... pensar en ese remoto combustible".

Menos de una generación más tarde, los ingleses recurrieron a la hulla, bajo la presión del elevado precio de la madera. A principios del siglo XVII se estimó que era imperativo que el gobierno pusiera empeño en detener la desforestación, porque la escasez de madera para la construcción naval parecía amenazar la existencia de Gran Bretaña. En una proclama real de 1615 se lamenta la antigua riqueza de "Leña y Madera", el tipo de



EL CARBON ERA EL PRINCIPAL COMBUSTIBLE de Gran Bretaña en las postrimerías del siglo XVII. Los cargadores, como los que aparecen en esta litografía de 1805, llevaban el carbón destinado a hogares e industrias del país y del extranjero. Al fondo, gabarras carboneras.

maderos que "no sólo son grandes y largos y voluminosos, sino que tienen también dureza y corazón, por cuanto no están expuestos a henderse o rajarse, y, por tanto, excelentes para su uso en barcos, como si Dios Todopoderoso, que había ordenado que esta Nación fuese poderosa por el Mar y la navegación, hubiera en su providencia dotado a la misma con el material principal conducente a ello." Hacia mediados del siglo XVII el carbón había demostrado ser tan útil, y tan difundido era su uso, que los ingleses habían llegado a convertir la necesidad en virtud. Se reconciliaron con la decepcionante incapacidad de sus exploradores para hallar fuentes de metales preciosos, y de sus mineros para hallar una buena cantidad de ellos en la misma Gran Bretaña.

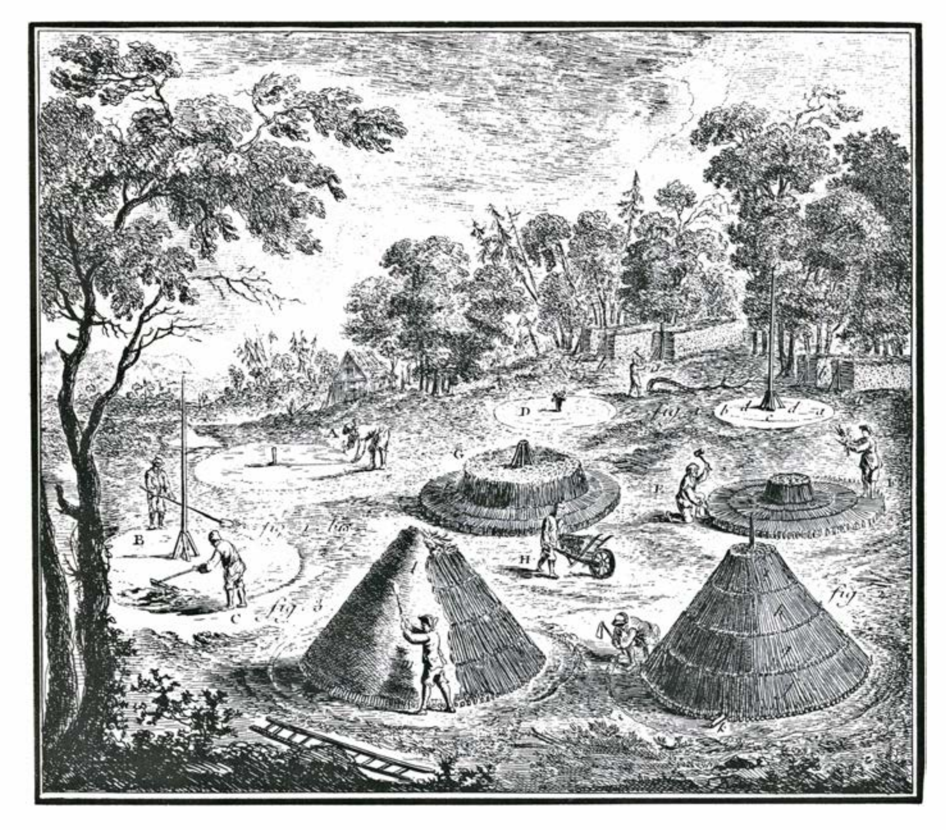
A pesar de los humos y vapores del carbón y del desagrado que provocaba, por doquiera, por la época de la guerra civil (1642-51), los londinenses dependían del embarque costero de carbón para defenderse del frío. En 1651, el anónimo autor de *News from Newcastle* escribía versos alabando el nuevo combustible "¡Inglaterra es un Mundo perfecto! ¡Tiene Indias también!/¡Corregid vuestros Mapas: New-Castle es Perú.../Que triunfe el malvado español, hasta que se diga/Que nuestro holliniento mineral purifica su oro".

Aún antes, como lo deja en claro la vista del crecimiento demográfico de la *Description of Britain* (1577) de William Harrison y una petición que los cerveceros ingleses dirigieron a Sir Francis Walsingham, secretario de estado de la reina Isabel (1578), el carbón estaba adquiriendo un nuevo e importante papel en la calefacción doméstica e industrial. Los archivos que aún quedan de los funcionarios de aduanas de Newcastle-on-Tyne (y registros subsiguientes de otras ciudades) revelan un crecimiento continuo y rápido de los embarques de carbón entre 1550 y 1700, primero desde Newcastle-on-Tyne y luego desde otros puertos. Estos registros sugieren que los embarques costeros aumentaron al menos por un factor de veinte entre los años citados. Las importaciones londinenses por vía marítima aumentaron con mayor celeridad, probablemente más de treinta veces su monto original, lo que no es de sorprender en vista del crecimiento demográfico de la ciudad en igual período. Lord Buckhurst, que llegó a ser tesorero de la reina Isabel a fines del siglo XVI, exigió a los funcionarios de aduana, durante la última década de ese siglo, que determinaran "la tasa de aumento" de los embarques de carbón, introduciendo de ese modo un nuevo concepto en los asuntos humanos. Los cálculos sobre los que insistía Buckhurst indicaban que podía contarse con que los impuestos proporcionaran una fuente de ingresos creciente, de modo que en 1599 y 1600 se impusieron impuestos sobre los embarques de carbón.

Las subidas más impresionantes en la tasa de aumento de la producción de carbón se produjeron en la segunda mitad del siglo XVI y a principios del siglo XVII. De hecho, la tasa de aumento del volumen de carbón extraído entre 1566 y 1606 quizá superó incluso la tasa de aumento (calculada a partir de estadísticas menos incompletas) del volumen extraído durante la primera parte del siglo XIX, es decir, en el clímax de la Revolución Industrial inglesa. Las cantidades efectivas de ese rápido desarrollo de la producción hullera tal vez parezcan insignificantes hoy en día, pero lo que es preciso recrear es la perspectiva de los isabelinos y de sus sucesores inmediatos, a quienes esa ampliación de las disponibilidades de carbón debió haberles parecido extraordinariamente rápida.

La hulla no sólo fue una fuente de energía, sino también un acicate para el desarrollo tecnológico. A la mayoría de los productos que podían fabricarse con fogatas abiertas de madera los dañaba el contacto con los humos del carbón. John A. Harris ha comentado que, a consecuencia de eso, "fueron escasas las ocasiones en que la adopción de la hulla no conllevó una significativa alteración de los procesos industriales". En efecto, los progresos tecnológicos de la Revolución Industrial fueron en gran medida la culminación del período innovador conexo al cambiar de la madera al carbón.

Fue preciso desarrollar nuevos métodos de caldeo, en los cuales los materiales por calentar se protegían de un contacto directo con las brasas y los gases de la combustión. De otro modo habría sido necesario convertir la hulla en coque, purgándola así de sus propiedades perjudiciales. Después de 1610, aproximadamente, el vidrio empezó a fabricarse con combustible mineral en una variante del horno de reververo, sistema que después desempeñó un importante papel en el crecimiento de otras industrias principales. En este tipo de horno, un techo abovedado refleja el calor de las brasas sobre el material por caldear, impidiendo así la contaminación de éste con las substancias procedentes del combustible. El carbonato potásico y la are-

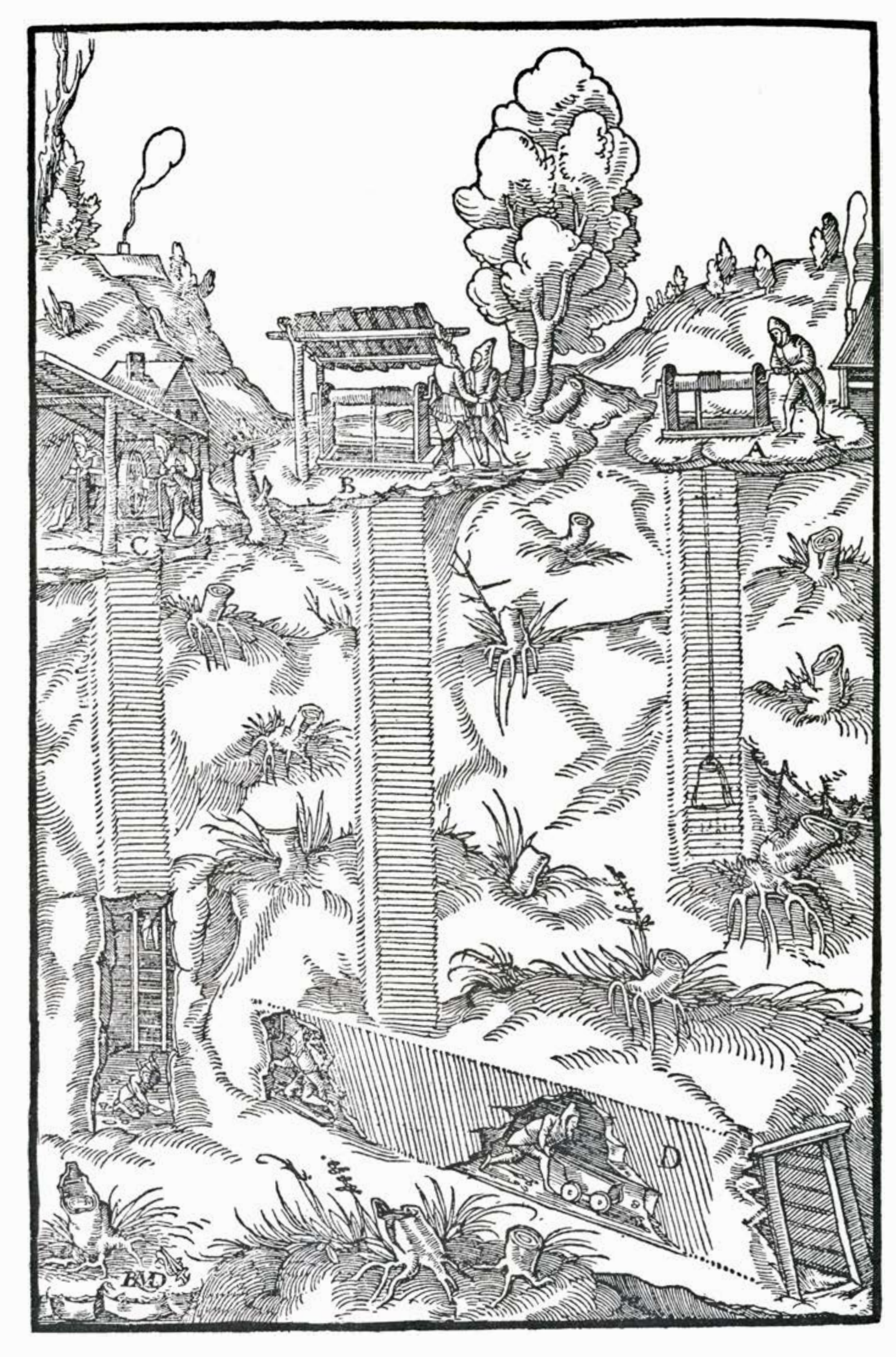


ANTES DEL ADVENIMIENTO DE LA HULLA, la madera era la principal fuente de energía térmica en Europa. El viento, los animales y el agua corriente proporcionaban la fuerza para la industria. A menudo, era preciso convertir la madera en carbón quemándola parcialmente en hornos como los ilustrados aquí. La madera se aplicaba en haces, se cubría con tierra y polvo de carbón y luego se le prendía fuego. La cubierta mantenía un mínimo de combustión, de manera que el producto final era carbón de madera y no ceniza. Para ciertos procesos fabriles se prefería el carbón a la leña, porque es casi carbón puro y, en cuanto entra en combustión, produce más calor. (*Encyclopédie, ou Dictionnaire Raisonné des Sciences, des Arts et des Métiers,* **de Diderot.**)

na que habían de fundirse para formar el vidrio se encerraban en un crisol de arcilla para protegerlos aún más de los humos. Al igual que el horno de reverbero, el crisol se empleó más tarde en otros muchos procesos fabriles.

Después de 1610, y en unas décadas, una nueva tecnología introdujo la hulla en muchos tipos de manufactura. El proceso de cementación para convertir hierro forjado en acero con carbón se introdujo entre 1612 y 1620. Hacia 1618, el embajador veneciano en Londres describió un método para cocer ladrillos en hogueras de hulla, en palabras que muestran que los italianos ya no estaban dispuestos a mirar en menos este "remoto combustible", como había recomendado Biringuccio. Antes de la guerra civil inglesa de la década de 1640, se introdujo el coque para el secado de la malta en la industria cervecera, que se había ampliado rápidamente durante la mayor parte del siglo XVI con la difusión del cultivo del lúpulo, copiado de los holandeses.

Una de las aplicaciones más importantes del carbón, después de la restauración de la monarquía inglesa en 1660, fue su uso en la adaptación del horno de reverbero para la fundición de metales no ferrosos. Esta innovación de la década de 1680-89 hizo posible fundir los minerales ingleses de plomo, cobre y estaño con carbón. A fines del siglo XVII sólo la producción de arrabio y hierro en barras seguía dependiendo de la madera. Aunque el problema no se solucionó por completo antes de 1784, se dio un paso importante para resolverlo en 1709, cuando Abraham Darby el viejo introdujo el coque en su alto horno, en Shropshire. En este tipo de horno, el combustible y el mineral están en contacto. El problema existente en el proceso de Darby consistía en que producía un tipo de arrabio que, a diferencia del elaborado con fuego de madera, no podía convertirse en hierro forjado. En 1784, Henry Cort inventó el proceso de pudelado, en el cual el arrabio (incluso el procedente de un alto horno) se vuelve a fundir y se manipula en un horno de reverbero caldeado por hulla, para producir hierro forjado. Hasta el invento de Cort, la fabricación de hierro siguió dependiendo en gran medida del carbón de leña. Así, aunque la producción inglesa de hierro se había multiplicado varias veces entre 1540 y cerca de 1620, este crecimiento se había visto frenado por la escasez de madera para hacer carbón después de este último año. Pero a partir de la década 1620-29, un aumento de las importaciones de hierro, especialmente desde Suecia, posibilitó un crecimiento lento y continuo de la fabricación de productos acabados de



LA CRISIS MADERERA del siglo XVI coincidió con una actitud distinta hacia la minería. Hasta la Edad Media, se había considerado en general la minería como una afrenta a la naturaleza. Pero en *De re metallica,* publicado en 1556, Georgius Agricola expresó un nuevo respeto por la minería. Esta cuidadosa relación de la metalurgia y la minería ofrece una buena imagen de esas industrias por la época en que fue necesario por primera vez aumentar la producción hullera. En esta ilustración de dicho libro, se ha cortado un túnel, *D,* en un cerro y se han excavado tres socavones desde arriba. Aunque la minería se facilitaba cuando un socavón conectaba con el túnel, no todos los socavones tenían esa finalidad. En este caso, el socavón *A* se explotará sólo desde la superficie. El socavón *B* conecta con el túnel, y el túnel pronto conectará con el socavón *C.* El material se izaba verticalmente de un socavón mediante un cabrestante, que solía taparse con un cobertizo para evitar la entrada de la lluvia. Agricola señalaba, a modo de consejo, que convenía construir un edificio aparte como vivienda, porque "a veces caen en los socavones muchachos y otros seres vivos".

hierro, que ya se hacían mediante procesos en que se usaba combustible mineral.

Samuel Eliot Morison hizo notar, acerca de las innovaciones en la construcción naval y en la navegación, que existe siempre "un lapso entre la invención de un dispositivo y el persuadir a los propietarios a suministrarlo o a los marinos a usarlo". Lo mismo puede decirse acerca de la difusión de los inventos relacionados con la introducción de la hulla en Inglaterra después de 1550. El hacer eficientes los nuevos métodos basados en el carbón llevó bastante tiempo. Por

ejemplo, en la fabricación de ladrillos (como también en el horneado de pipas de barro), cuando se introdujo el horno caldeado con carbón, había mucho desperdicio por rotura. Pero antes de terminar el siglo XVII, ya se perdían pocos ladrillos en ese sistema.

Conforme fue haciéndose patente que la hulla podía significar una producción más barata y eficaz, aumentaba el número de industrias que la adoptaban como combustible. Antes de 1700, en la creciente industria textil británica, en la cual procesos tales como el tratamiento con vapor y el teñido exigían grandes cantidades de combustible, éste solía ser el carbón. Antes de dicha fecha, las crecientes manufacturas de sal, alúmina, caparrosa (vitriolo verde, o sulfato ferroso), salitre, pólvora, almidón y velas, dependían del carbón. También se lo usaba para entonces muchísimo en la preparación de conservas, de vinagre y whisky escocés, y en las cervecerías, jabonerías y refinación de azúcar. Un visitante francés, que estudiaba la tecnología de los Midlands en 1738-39, informaba que los nuevos hornos caldeados con carbón (hechos de ladrillos cocidos al carbón) habían producido un fertilizante de cal de tal calidad que se había triplicado el rendimiento de la tierra laborable; consideraba al carbón como "el alma de las manufacturas inglesas".

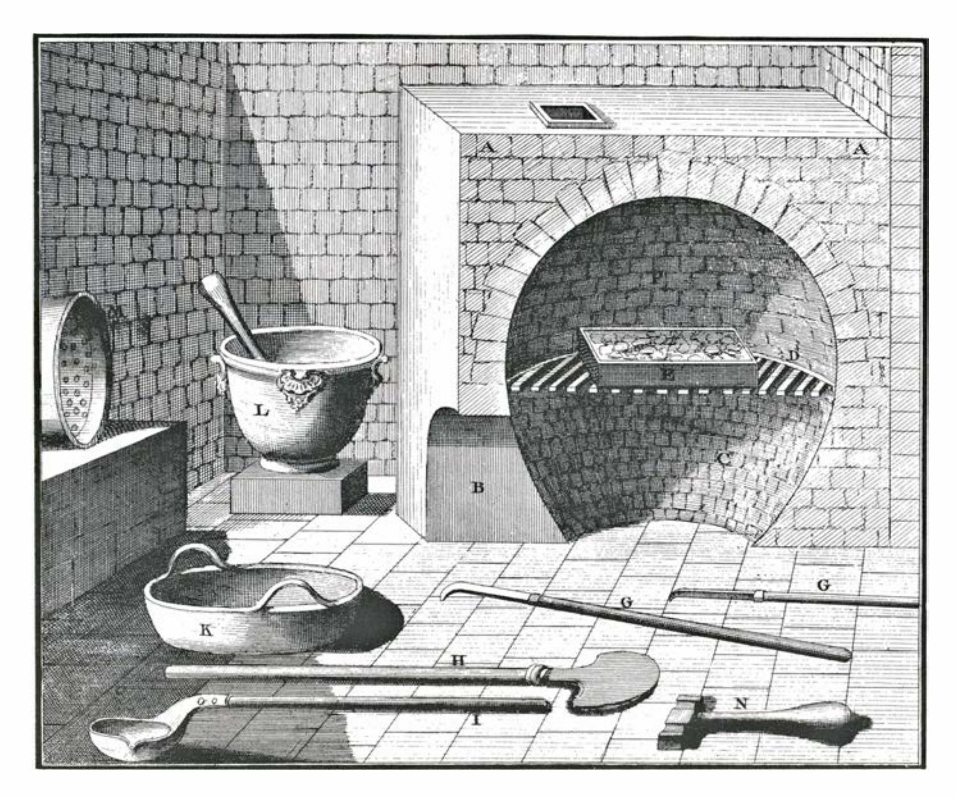
La difusión del uso del carbón en los hogares ingleses, que comenzó en los inicios de la época isabelina, fue continua durante todo el siglo XVII. No fue éste el único cambio residencial originado por la adopción del combustible mineral en Inglaterra. Bajo la reina Isabel y sus sucesores Estuardo, el reino fue reconstruido extensamente. Estructuras de ladrillo y de piedra (con mortero hecho de caliza calcinada con hulla) iban reemplazando las de madera. Ventanas de vidrio (producido en hornos a carbón) se instalaban en los edificios para retener el calor procedente de los nuevos hogares a carbón (que tenían parrillas de hierro y chimeneas de ladrillo fabricados con carbón). A pesar de su hollín y su hedor, el carbón había traído un nuevo bienestar al clima húmedo y frío de Inglaterra. Ya en 1651, el autor de *News from Newcastle* observaba ¡que los sacos de carbón habían acrecentado las alegrías de la intimidad!

El carbón se había incorporado tan satisfactoriamente a la tecnología y economía inglesas, que durante las últimas cuatro décadas del siglo XVII los precios de la madera dejaron de subir. Hace unos años aventuré a ojo de buen cubero una cifra de tres millones de toneladas para la producción hullera anual de Inglaterra en la década de 1690. En opinión de Harris, esa cantidad "quizá resulte ser, con el tiempo, demasiado prudente y no excesiva". Parece que en aquel entonces se estaba llevando a cabo cuatro veces más calefacción que la habida precedentemente con madera. Nunca antes un país de importancia había llegado a depender de sus recursos subterráneos para el grueso de su abastecimiento de combustible.

Aunque la explotación del carbón había resuelto en gran parte la escasez de combustible antes de 1700, había todavía escasez de madera porque habían aumentado otras demandas de madera. En 1618, un viajero venido de Londres describía su tiempo como una "época ruidosa, retintineante, estrepitosa", y anotaba que "el Mundo marcha sobre ruedas de madera". Para la construcción de barcos y vehículos de tracción animal, necesarios para transportar personas y mercancías a través de agua y tierra, se requerían grandes cantidades de madera. Además, aunque hubo cierta repoblación forestal durante el siglo XVII, se talaba más y más bosques para dejar paso a cultivos y ganadería. Asimismo, se desbrozaba zonas más pequeñas para la ampliación de la minería, especialmente de las hulleras. Sencillamente, los bosques de Inglaterra no podían seguir el ritmo de la demanda de madera de la isla.

Los ingleses se vieron obligados a suplementar su abastecimiento interno con importaciones, principalmente traídas de las colonias norteamericanas y de la región báltica. (En *La riqueza de las naciones*, publicado en 1776, Adam Smith observaba que en su Edimburgo natal "no había quizá ni un solo palo de madera escocesa".) Parte de las importaciones se pagaba mediante las crecientes exportaciones de carbón y, probablemente en mayor proporción, mediante las crecientes exportaciones de textiles fabricados en diverso grado con carbón combustible. Este comercio exterior, y aun más el comercio de cabotaje en rápida expansión, había traído por consecuencia, ya en el siglo XVII, el desarrollo de una gran flota mercante inglesa. Nuevos barcos carboneros se proyectaban para llevar más hulla con menos tripulación, y el comercio hullero de cabotaje se consideraba como la principal escuela para marinos, factor importante en el surgimiento de Inglaterra como potencia marítima.

Sin embargo, el carbón, en ciertos casos, hizo a Inglaterra más independiente respecto de mercancías importadas, como, por ejemplo, de la sal. Según explica Robert Multhauf en su libro de próxima aparición: *Neptune's Gift: A History of*



EL HORNO DE REVERBERO hizo posible emplear la hulla a pesar de los humos y llamas reactivos del combustible. El techo abovedado del horno refleja el calor de la combustión en el material por calentar. Cuando el combustible que se quema es hulla, la disposición impide que las sustancias de los humos del carbón contaminen el material. Esta ilustración de un horno de reverbero para recocido está tomado del capítulo sobre amonedado de la *Encyclopédie* de Diderot. Los troqueles, como el mostrado dentro del horno, habían de recocerse antes de poder estampar las monedas.

Common Salt (Un don de Neptuno: Historia de la sal común), esta mercancía era indispensable en Europa durante los siglos XVI y XVII. En Inglaterra, en donde los alimentos de origen marino iban a ocupar un lugar más importante en una dieta cada vez más abundante, la sal era indispensable para conservar el pescado. En el sur y el poniente de Francia, la sal se obtenía dejando que el sol hiciera evaporarse el agua de mar en charcas o estanques de poco fondo, pero este método era poco práctico para el clima británico. A comienzos del siglo XVI, dos tercios de la sal consumida en Inglaterra había de importarse, en su mayor parte de Francia. El cambio casi total de combustible en Inglaterra, de la madera al carbón, modificó radicalmente esa situación. A finales del siglo XVII, se quemaban unas 300.000 toneladas, es decir, casi el 10 por ciento del carbón extraído anualmente, para evaporar agua y producir así sal en Inglaterra y Escocia. El resultado fue que el país llegó a ser prácticamente autosuficiente en cuanto a sal.

La conversión a un nuevo tipo de combustible podría haber tenido un efecto menor en la economía británica si el país hubiese estado dotado escasa, o siquiera moderadamente, de carbón. Pero antes de terminar el siglo XVII había llegado a ser evidente que poseía enormes reservas hulleras. Un producto de la tecnología inspirada en el carbón proporcionó nuevas y tranquilizadoras informaciones. El dispositivo, denominado varilla de sonda, se introdujo a principios del siglo XVII. Las primeras prospecciones con varilla de sonda fueron inexactas, pero antes de finalizar ese siglo los expertos en minería podían determinar el espesor y calidad de las capas carboníferas sin hacer socavones. Las varillas de sonda se habían convertido en herramientas fiables y habían revelado una nueva tierra de promisión hallada bajo la superficie y aún bajo los mares vecinos. Se veía que bajo gran parte de la isla yacía carbón. Este tesoro de recursos energéticos comenzó a ejercer un empuje en la dirección de una producción en cantidad que no había tenido par en la historia precedente.

Hasta mediados del siglo XIX, después que se iniciara una aceleración nunca vista de la tasa de crecimiento de la producción, y sólo tras la publicación, en 1865, de *The Coal Question,* obra de William Stanley Jevons, algunos no se percataron de que los yacimientos de carbón eran agotables. Para entonces, se conocían recursos de petróleo y de gas natural fuera de Inglaterra, aunque ninguno de ambos se explotara hasta más avanzado el siglo XIX. Hubo que esperar hasta la década de 1920 para que algunas personas comenzaran a darse cuenta de que las disponibilidades de todos los combustibles fósiles tenían límites definidos.

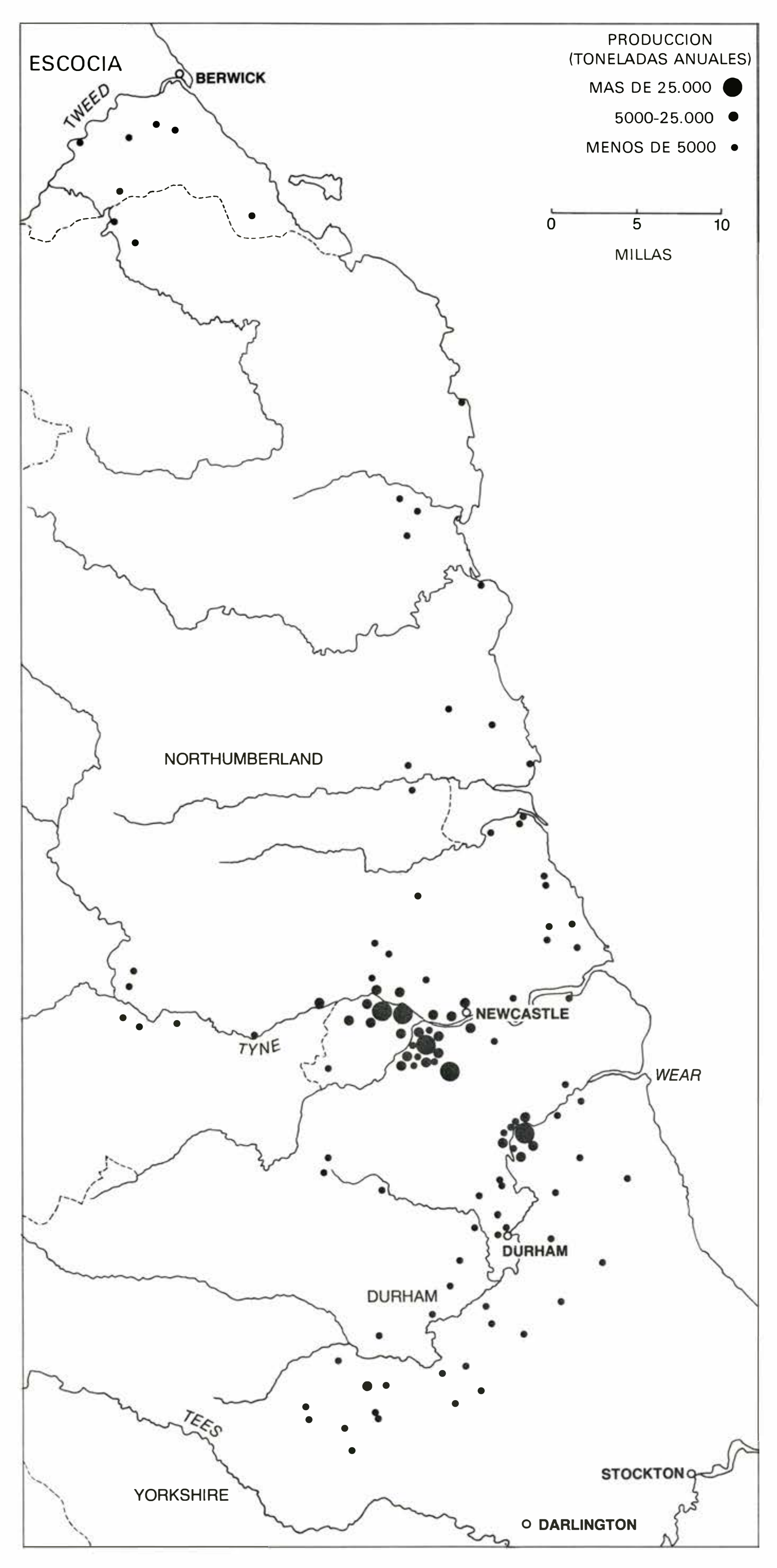
El cambio hacia el combustible fósil ocurrido en el siglo XVII llevó, después de 1780, a la explotación agresiva de las vastas reservas mundiales de mineral de hierro. Sin el advenimiento de la primera economía energizada por carbón, la edad del hierro y el acero quizá no se hubiese desarrollado nunca. El cambio hacia la hulla, iniciado en la Inglaterra isabelina, tuvo aún otras consecuencias al dar a luz la era moderna y mecanizada. El empleo de las máquinas a vapor y de los viajes por tren fueron también vitales para el nacimiento de esa era. Los intentos por construir motores a vapor y de introducir vías de carriles con vagones arrastrados por caballos se remontan en Inglaterra al reinado de Jacobo I, pero no fue sino hasta 1712 cuando Thomas Newcomen instaló en una mina de carbón de Staffordshire una máquina a vapor que funcionaba realmente. Las necesidades de la minería del carbón y de su transporte indujeron en medida no pequeña la creación de la máquina de vapor de agua y del ferrocarril. El clima húmedo de Inglaterra agravaba el problema de los perjuicios causados por el agua en el creciente número de hulleras. La potencia de los caballos (que comían costoso forraje) y del agua corriente (que requería capital para presas y ruedas hidráulicas de arcaduces) estuvo disminuyendo los beneficios procedentes de la minería del carbón a lo largo de todo el siglo XVII. La imperiosa necesidad de sistemas más eficientes de avenamiento en las hulleras inglesas en esa primera época desempeñó un importante papel en la creación de la máquina de Newcomen. Una vez que se inventaron estos "ingenios de fuego", según lo han demostrado John S. Allen y Alan Smith, se difundieron pronto por las islas británicas entre 1712 y 1730, aproximadamente.

Debe recordarse que el carbón parece haber sido usado como combustible para muchas finalidades industriales en China, en la era Sung, en los siglos X y XI. El episodio fue olvidado en gran parte. Obviamente no condujo, como en Europa, mucho más tarde, a una revolución industrial.

Los estudios realizados en estos últimos 50 años, desde que publiqué mis dos volúmenes sobre el carbón, han mostrado que el período de la crisis energética de Inglaterra –fines del siglo XVI y todo el XVII– fue también el período de lo que ha dado en llamarse la revolución científica. La revolución del pensamiento que hizo nacer la ciencia moderna fue un factor aún más importante que la hulla en el establecimiento de la era mecanizada. Hacia las décadas segunda y tercera del siglo XVII, los europeos se estaban percatando del inmenso desarrollo de la producción que prometía el desarrollo del nuevo combustible. En esas dos décadas escribió Francis Bacon *La nueva Atlántida* (1627) y René Descartes su *Discurso del método* (1637). La isla que Bacon imaginaba en un mar lejano estaba provista de una gran institución de investigación científica, que presidía sobre el destino humano, y confiaba en que una nueva abundancia, que el desarrollo del conocimiento científico haría posible, resolvería tanto los problemas intelectuales y morales como los económicos. Y Descartes, en su *Discurso,* no se mostraba menos esperanzado. En forma aún más específica que Bacon, previó una mayor producción, un trabajo más liviano y una más larga vida para los seres humanos por doquiera. Incluso en aquella época se hablaba de vehículos aéreos, submarinos, explosivos devastadores y viajes a la Luna. Fue la revolución científica de fines del siglo XVI, y del siglo XVII, conjuntamente con las transformaciones económicas suscitadas por la introducción del carbón, la que dio nacimiento al mundo industrial en que vivimos.

En Inglaterra, el planteamiento y resolución de la crisis energética (1550 a 1700) se caracterizaron por una creciente rentabilidad del trabajo en todos los tipos de producción. El estadista e historiador inglés Lord Clarendon (1609-1674) se estaba refiriendo a esa prosperidad cuando escribió que durante el período anterior a los años 1640, los ingleses "disfrutaron... la más plena medida de la felicidad con que pueblo alguno haya sido antes bendecido por tan largo tiempo; para maravilla y envidia de todos los confines de la Cristiandad". Clarendon no basó su afirmación en lo que se considerarían hoy unas estadísticas fiables, y en verdad no es posible procurar tales estadísticas. Sin embargo, las nuevas informaciones sobre el desarrollo de la economía inglesa contenidas en libros recientes de Eric Kerridge y del autor de este artículo, respaldan la afirmación de Clarendon. Inglaterra, que estaba retrasada respecto del resto de Europa en muchos sectores económicos durante la Edad media tardía, probablemente aventajaba a los demás países europeos en producción agrícola e industrial per capita hacia 1700.

Kerridge ha demostrado que mientras se había pensado que la agricultura bri-



EL MAPA DE LOS CONDADOS DE DURHAM Y NORTHUMBERLAND muestra los emplazamientos aproximados de hulleras en 1635. Probablemente había más hulleras que las indicadas en el mapa. (Ilustración adaptada de un mapa de *The Rise of the British Coal Industry,* del autor.)

tánica se había transformado en el siglo XVIII e inicios del XIX, la transformación se efectuó de hecho entre fines de la Edad Media y el término del siglo XVII. En este período, escribe, "el aumento de los rendimientos (en el campo) debe haber sido enorme. Los rendimientos de cereales y pastos se cuadruplicaron, aproximadamente, y los rendimientos de los barbechos (tierras no labradas anteriormente) aumentaron hasta ser irreconocibles. En conjunto, es difícil negarse a la conclusión de que, en promedio, los rendimientos llegaron hasta decuplicarse y quintuplicarse".

La mayoría de quienes estudiaron, entre 1920 y 1940, el advenimiento del capitalismo (y yo incluido) aceptaba las explicaciones de Karl Marx, Sir William Ashley, Max Weber, Henri Hauser y otros. Los trabajos de estos eruditos sugerían que el nacimiento del capitalismo y del "espíritu capitalista" era el factor principal conducente al aumento abrumador de la producción de bienes y servicios en el siglo XIX. Actualmente, pienso que un factor más importante fue una fe creciente en el progreso cuantitativo, en la multiplicación de la producción.

Hacia fines del siglo XVI, empezó a enfocarse una nueva atención sobre los conceptos de cantidad. Los efectos de este nuevo interés pueden observarse en las mediciones más exactas empleadas en las ciencias naturales en vías de desarrollo y en el reemplazo del calendario juliano por el gregoriano, muchísimo más exacto. Se reflejó también en una serie de invenciones concebidas para acelerar los cálculos numéricos, una realizada por Galileo (1564-1642), otra por el matemático holandés Simon Stevin (1548-1620) y dos más por el hacendado escocés John Napier (1550-1617), creador de los logaritmos. Unas matemáticas complejas –el cálculo– se desarrollaron primero en Francia después de 1620 y más plenamente, avanzado ya el siglo XVII, por obra de Newton y de Leibniz. La idea de unas tasas de crecimiento, introducida durante la época isabelina, trajo una novedosa precisión a los estudios económicos. La nueva perspectiva destacaba el probable valor de metas cuantitativas para la humanidad. La transformación de los objetivos industriales constituyó un importante avance hacia un mundo industrializado.

En 1697, un inglés, llamado James Puckle, escribió: "Nuestros artesanos son universalmente considerados los mejores de la tierra en cuanto a mejoramientos". Esto era sin duda verdadero en la manufactura que exigía eficiencia y producción en cantidad. Pero se requiere

efectuar una valoración diferente del estado de las artes y de las artesanías de lujo en Inglaterra después del cambio de combustible. En la conjunción de los siglos XVII y XVIII, Europa estaba ansiosa por aprender de Gran Bretaña unos métodos de producción más eficaces, pero los ingleses estaban igualmente ávidos de aprender las maneras de confeccionar productos y ambientes hermosos de parte de italianos, franceses y holandeses. (En ninguna parte de Europa era más notable la búsqueda de la belleza y la armonía en edificios y mobiliarios que en los Países Bajos de Rembrandt y Vermeer.) Harris ha expuesto que en el siglo XVIII los ingleses, a pesar de sus aspiraciones a la gran moda, se veían en serios problemas para copiar los métodos de fabricación de cristales de alta calidad que empleaban los franceses en Saint-Gobain. En Gran Bretaña, el gran desarrollo de la industria del carbón había debilitado la posición de la artesanía y del arte como alma y corazón de la producción.

Además, el auge de la minería del carbón había ensombrecido a los trabajadores relacionados con el carbón. Los mineros del carbón y transportistas, manchados por el negro mineral, eran a menudo unos parias. Se les veía como hombres negros, y en el siglo XVII, cuando a los hombres epidérmicamente negros se les despachaba como esclavos de Africa a América, a los trabajadores del carbón se les estaba sometiendo a una nueva forma de esclavitud en las hulleras escocesas y en las bateas de sal caldeadas a hulla.

Conforme el carbón se difundió desde Inglaterra al resto de Europa, hacia fines del siglo XVIII, y de allí en adelante, menguó la preocupación por la belleza en las manufacturas y por el medio ambiente humano. A lo largo de la historia, esta especie de dedicación a la belleza ha sido importante en la fijación de topes razonables para el desarrollo económico. El advenimiento del carbón parece haber disminuido esa dedicación. La explotación de los recursos de la tierra ha violado a menudo los límites del buen gusto. Usar al máximo esos recursos exige no sólo ingenio, sino también moderación. En la actualidad, la dependencia del hombre respecto de los combustibles fósiles es tan problemática como lo fue su dependencia de la madera hace unos 400 años. La mayor esperanza para la explotación fructuosa de los recursos de combustibles quizá resida en una renovación y amplificación de las normas de belleza. Para que la humanidad siga adelante, hacer la historia debe convertirse en un arte, esto es, en una búsqueda de la belleza.

La estructura de los conceptos científicos

La asombrosa variedad de los conceptos usados en la ciencia es reducible a sólo tres tipos distintos. El estudio de la estructura matemática de cada uno de ellos arroja nueva luz sobre nuestra comprensión de la conceptualización científica

Jesús Mosterín

El mundo nos bombardea continuamente con todo tipo de radiaciones, roces, mensajes. Y nuestro aparato sensorial selecciona y procesa esa información bruta que nos llega del mundo. Si tuviéramos otro aparato sensorial diferente del que tenemos, percibiríamos el mundo de distinto modo. Si nuestra retina fuese sensible a otro intervalo distinto del espectro electromagnético, veríamos un paisaje infrarrojo o ultravioleta muy distinto del que vemos. Si nuestros oídos fueran sensibles a otras frecuencias, escucharíamos un mundo ahora inaudito para nosotros. Y si tuviéramos sentidos sensibles a la radiactividad o al magnetismo, percibiríamos el mundo de un modo ahora inimaginable. Esto no significa que nuestros sentidos inventen el mundo ni que nuestras percepciones no sean objetivas. Tan objetiva es una foto en blanco y negro como una foto en color y como una radiografía. Pero nuestro aparato sensorial condiciona nuestra percepción del mundo y determina las pautas en las que ésta es posible. El mundo percibido es la resultante de al menos dos factores: nuestro aparato sensorial y el mundo exterior.

De igual modo, lo que pensemos y digamos del mundo no depende sólo de él, sino también de nuestro sistema conceptual, que selecciona, condiciona y determina los aspectos del mundo que tenemos en cuenta, en los que pensamos y de los que hablamos. El mundo pensado es también la resultante de al menos dos factores: nuestro sistema conceptual y el mundo real.

En nuestra actividad científica tenemos que partir de nuestro aparato sensorial y del sistema conceptual plasmado en nuestro lenguaje ordinario o común. Pero difícilmente podría ponerse en marcha la empresa científica si no nos fuera posible trascender las limitaciones de nuestro aparato sensorial y conceptual. Mediante instrumentos materiales apropiados, que son como extensiones de nuestros sentidos –telescopios, microscopios, cámaras fotográficas y de cine, balanzas, voltímetros, cuentarrevoluciones, veletas, brújulas, barómetros, magnetófonos, antenas de radio, etcétera– podemos discriminar mucho más finamente que con nuestros sentidos y podemos captar mensajes y radiaciones inasequibles a nuestro aparato sensorial. De igual modo, podemos extender y precisar nuestro sistema conceptual introduciendo conceptos más precisos y de mayor alcance que los del lenguaje ordinario, conceptos científicos que nos permiten describir hechos y formular hipótesis con una precisión y universalidad crecientes.

El progreso de la ciencia no siempre consiste en el aumento del número de verdades que conocemos. La noción de verdad es relativa a la de enunciado, y ésta a la de concepto. Qué verdades haya depende de qué conceptos empleemos. Y muchas veces el progreso de la ciencia consiste no en un aumento del número de verdades expresadas con un sistema conceptual dado, sino en el cambio del sistema conceptual, en su ampliación o extensión o en su sustitución por otro.

El mundo no está estructurado de por sí de un modo unívoco. Somos nosotros los que lo estructuramos al proyectar sobre él nuestros conceptos. Así, propiedades como la temperatura o la inteligencia no son intrínsecamente cualitativas o cuantitativas, sino que ese carácter sólo está en los conceptos que empleemos para hablar de ellas. Sin embargo, una vez introducidos ciertos conceptos de un determinado modo, ya no podremos usarlos a nuestro antojo, sino sólo siguiendo los perfiles que la realidad adopte al proyectar sobre ella dichos conceptos.

El importante papel desempeñado por los conceptos en la teorización científica ha despertado el interés de los metodólogos y filósofos de la ciencia, que en las dos últimas décadas les han prestado una atención especial. Lo primero que salta a la vista es la gran variedad de los conceptos científicos. Unos –como *pez, fuerza* o *calor*– proceden del lenguaje ordinario o común, algunas de cuyas nociones intuitivas precisan; otros –como *ARN mensajero, fonema* o *entropía*– constituyen creaciones artificiales ligadas a nuevos descubrimientos o teorías. Pero unos y otros se articulan de mil modos distintos en el seno de múltiples y heteróclitas teorías. ¿Cómo hincar el diente en esta profusión de conceptos distintos? La investigación reciente ha mostrado que uno de los puntos de vista más fecundos para el estudio metacientífico de los conceptos es el de su estructura formal o matemática. De hecho, la profusa variedad de los conceptos científicos se reduce desde este punto de vista a unos pocos tipos básicos, fundamentalmente a tres: los conceptos clasificatorios, los comparativos y los métricos.

Clasificaciones: condiciones formales de adecuación

Un concepto clasificatorio sirve para referirnos a un grupo determinado de objetos o sucesos que tienen algo en común. Los sustantivos y adjetivos del lenguaje ordinario suelen corresponder a conceptos clasificatorios: hombre, mujer, árbol, camión, azul, puntiagudo, muerto. Algunos de los conceptos clasificatorios del lenguaje ordinario *-bicho, pájaro, enorme–* son demasiado vagos para poder ser incorporados en el lenguaje científico, pues no determinan unívocamente la clase de las cosas a las que se

aplican. Sin embargo otros, más precisos –como *urraca, olmo* o *hirviente*– pueden ser incorporados sin más trámite que el de la explicitación de las notas comunes a todos los objetos a los que se aplican. De todos modos, el repertorio de conceptos clasificatorios de un lenguaje natural determinado –sea el náuhatl o el inglés, el swahili o el italiano– es siempre muy limitado y claramente insuficiente para las necesidades de la ciencia. Así, cada pueblo suele disponer de conceptos de los animales y plantas visibles y frecuentes en la zona que habita, pero no de los seres vivos invisibles a simple vista o de los animales de otras partes del mundo. Por ello, las comunidades científicas se ven obligadas a introducir numerosos conceptos clasificatorios nuevos y artificiales en el lenguaje científico.

En la ciencia, los conceptos clasificatorios no suelen introducirse aisladamente, sino en conjuntos llamados clasificaciones. Para que una clasificación –o sistema de conceptos clasificatorios– sea aceptable ha de cumplir dos tipos de condiciones de adecuación. Por un lado, unas condiciones formales de adecuación, comunes a todas las ciencias, y, por otro, ciertas condiciones materiales de adecuación peculiares de la ciencia de que se trate.

En una de sus obras *(El idioma analítico de John Wilkins),* el escritor argentino Jorge Luis Borges cita una imaginaria enciclopedia china, según la cual "los animales se dividen en (a) pertenecientes al Emperador, (b) embalsamados, (c) amaestrados, (d) lechones, (e) sirenas, (f) fabulosos, (g) perros sueltos, (h) incluidos en esta clasificación, (i) que se agitan como locos, (j) innumerables, (k) dibujados con un pincel finísimo de pelo de camello, (l) etcétera, (m) que acaban de romper el jarrón, (n) que de lejos parecen moscas". Esta presunta clasificación nos choca y sorprende porque viola completamente las condiciones formales de adecuación que esperamos satisfaga. En efecto, aunque el ámbito de objetos a clasificar parece ser el de los animales, algunos de los conceptos no se refieren a animales (como los dibujados con un pincel o las sirenas), otros no se sabe a qué se refieren (etcétera), los mismos animales caen bajo varios de estos conceptos (pertenecientes al Emperador, amaestrados), hay animales que no caen bajo ninguno de esos conceptos, y así sucesivamente.

En general, cuando hablamos de una clasificación esperamos que esté perfectamente delimitado cuál sea el ámbito o dominio de individuos que vamos a clasificar, que a cada concepto clasificatorio corresponda al menos un individuo de ese ámbito, que ningún individuo caiga bajo dos conceptos clasificatorios distintos y que todo individuo del ámbito en cuestión caiga bajo alguno de los conceptos de la clasificación.

La extensión de un concepto es la clase de las cosas a las que ese concepto se aplica. Si identificamos los conceptos clasificatorios con sus extensiones, entonces podemos resumir las condiciones formales de adecuación de una clasificación diciendo que la clasificación debe constituir una partición, en el sentido matemático de este término.

Sea A una clase de objetos. Sean $B_1 ... B_n$ partes de A, es decir, clases de objetos de A. Decimos que $H = \{B_1 ... B_n\}$ es una *partición* de A si y sólo si se cumplen las siguientes tres condiciones: (1) ninguna de las clases $B_1 ... B_n$ es vacía, (2) no hay ningún elemento común a dos de ellas y (3) entre todas abarcan a todos los elementos de A.

Así, la clasificación de los mamíferos en órdenes constituye una partición del conjunto de los mamíferos.

Particiones y relaciones de equivalencia

El concepto de partición está estrechamente ligado al de relación de equivalencia. Como es bien sabido, una relación de equivalencia es una relación reflexiva, simétrica y transitiva. Es decir, una relación R es una *relación de equivalencia* en una clase A si y sólo si (1) todo individuo de A está en la relación R consigo mismo, (2) si x está en la relación R con y, entonces y también está en la relación R con x, y (3) siempre que u esté en la relación R con w y w con z,

$\neg\ ...$	:	no ...
$...\wedge...$	:	... y ...
$...\vee...$	:	... o ...
$...\Rightarrow...$	:	si ..., entonces ...
$...\Leftrightarrow...$	:	... si y sólo si ...
$\bigwedge x$	:	para todo x
$\bigvee x$	:	hay un x, tal que
$...\in...$	:	...es elemento de...
$...\notin...$	:	...no es elemento de...
$...\subset...$	:	...está incluido en...
$\emptyset$	:	conjunto vacío
$...\times...$	:	producto cartesiano de ... por ...
$\mathcal{P}...$	:	conjunto de las partes de ...
$...<...$	:	...es menor que...
$...=...$	:	...es igual a...
$...\neq...$	:	...es distinto de...

Sea $R \subset A \times A$

$$R \text{ es una relación de equivalencia en } A \Leftrightarrow \bigwedge x \in A : xRx \ \wedge \ \bigwedge xy(xRy \Rightarrow yRx) \ \wedge \ \bigwedge xyz(xRy \wedge yRz \Rightarrow xRz)$$

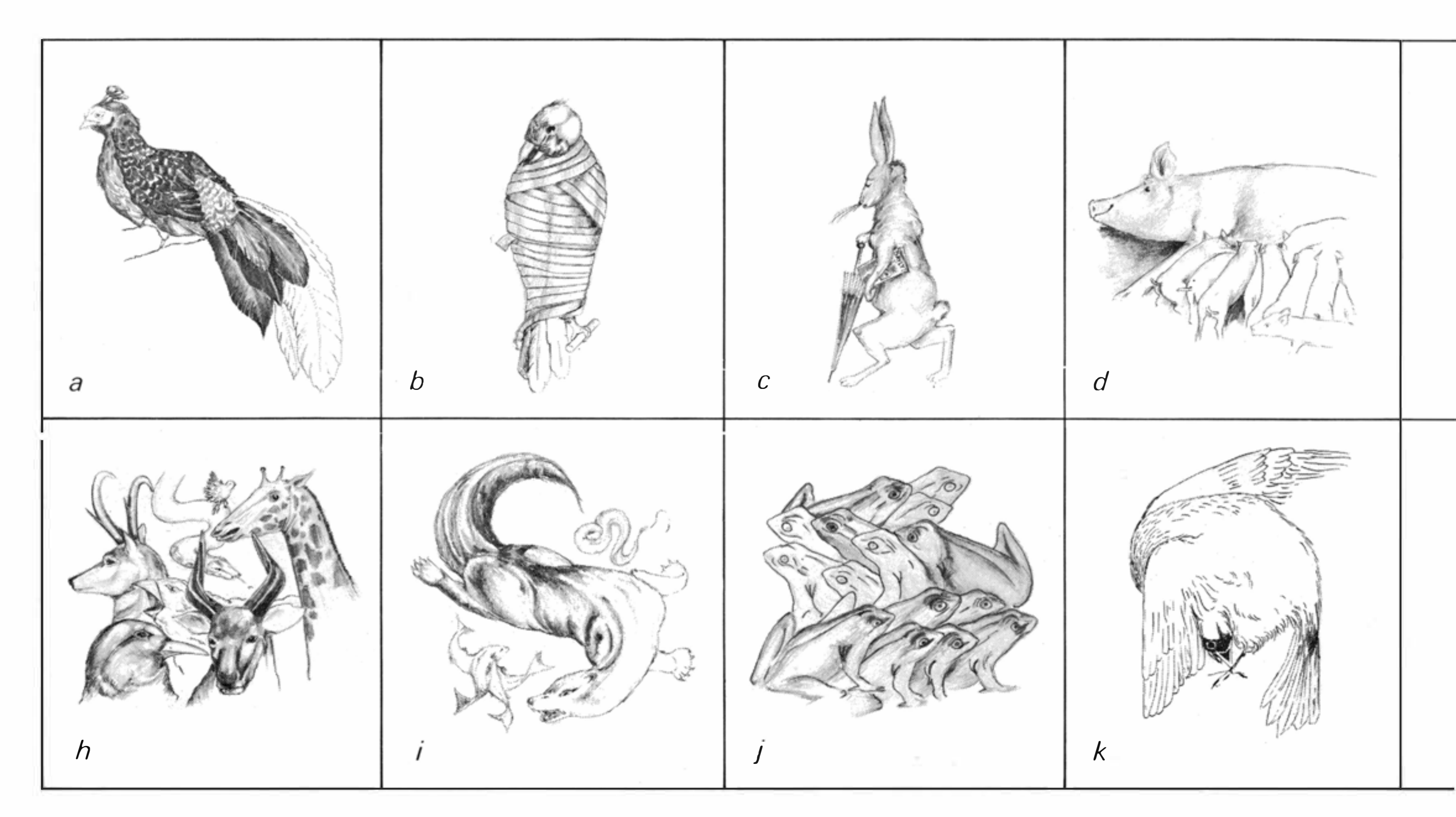
Sea $H \subset \mathcal{P}A$

$$H \text{ es una partición de } A \Leftrightarrow \emptyset \notin H \ \wedge \ \bigwedge xy \in H : (x \neq y \Rightarrow x \cap y = \emptyset) \ \wedge \ \bigcup H = A$$

Sean $K \subset A \times A$ y $P \subset A \times A$

$$\langle A, K, P \rangle \text{ es un sistema comparativo} \Leftrightarrow \bigwedge x \in A : xKx \ \wedge \ \bigwedge xy\,(xKy \Rightarrow yKx) \ \wedge \ \bigwedge xyz\,(xKy \wedge yKz \Rightarrow xKz) \ \wedge \ \bigwedge xyz\,(xPy \wedge yPz \Rightarrow xPz) \ \wedge \ \bigwedge xy\,(xKy \Rightarrow \neg\, xPy) \ \wedge \ \bigwedge xy\,(xKy \vee xPy \vee yPx)$$

UNA DEFINICION PRECISA Y CONCISA de las nociones de relación de equivalencia, partición y sistema comparativo puede obtenerse mediante el uso de signos lógicos y matemáticos apropiados –como los introducidos en la parte superior del presente cuadro– y de letras y variables, que expresan la generalidad de dichas nociones, aplicables en los más diversos contextos y situaciones.



CLASIFICACION FANTASTICA de los animales en la enciclopedia china imaginada por Borges, según la cual "los animales se dividen en (*a*) **pertenecientes al Emperador,** (*b*) **embalsamados,** (*c*) **amaestrados,** (*d*) **lechones,** (*e*) **sirenas,** (*f*) **fabulosos,** (*g*) **perros sueltos** (*h*) **incluidos en**

también u está en la relación R con z. Por ejemplo, la identidad es una relación de equivalencia, así como también lo es la de paisanaje entre humanos y la de congruencia entre triángulos.

Sea R una relación de equivalencia en A. Sea x un elemento de A. Mediante x_R designamos la clase de equivalencia de x respecto a R, es decir,

$$x_R = \{y \epsilon A | y\ R x\}$$

De aquí se sigue que las clases de equivalencia de x e y respecto a R serán la misma si y sólo si x está en la relación R con y.

Llamemos *conjunto cociente* de A respecto a R (en signos A/R) al conjunto de todas las clases de equivalencia de elementos de A respecto a R, es decir,

$$A/R = \{x_R | x \epsilon A\}$$

Por definición, ninguna de las clases de equivalencia de A/R es vacía, dos clases de equivalencia distintas de A/R no tienen elementos comunes (pues si los tuviesen serían la misma clase) y entre todas abarcan A entero, pues para cada elemento $x \epsilon A$, x está en x_R y, por tanto,

$$\bigcup_{x \epsilon A} x_R = A$$

Por tanto, si R es una relación de equivalencia en A, el conjunto cociente de A respecto a R es una partición de A.

A la inversa, toda partición da lugar a una relación de equivalencia. Sea $H = \{B_1, \ldots B_n\}$ una partición de A. Podemos definir la relación R_H entre elementos de A del siguiente modo: para cada dos elementos x y z de A, x está en la relación R con z si y sólo si tanto x como z están en la misma subclase B_i de A. Es decir,

$$R_H = \{\langle x, z \rangle \epsilon A^2 | \text{hay un i } (1 \leqslant i \leqslant n) \text{ tal que } x \epsilon B_i \text{ y } z \epsilon B_i\}$$

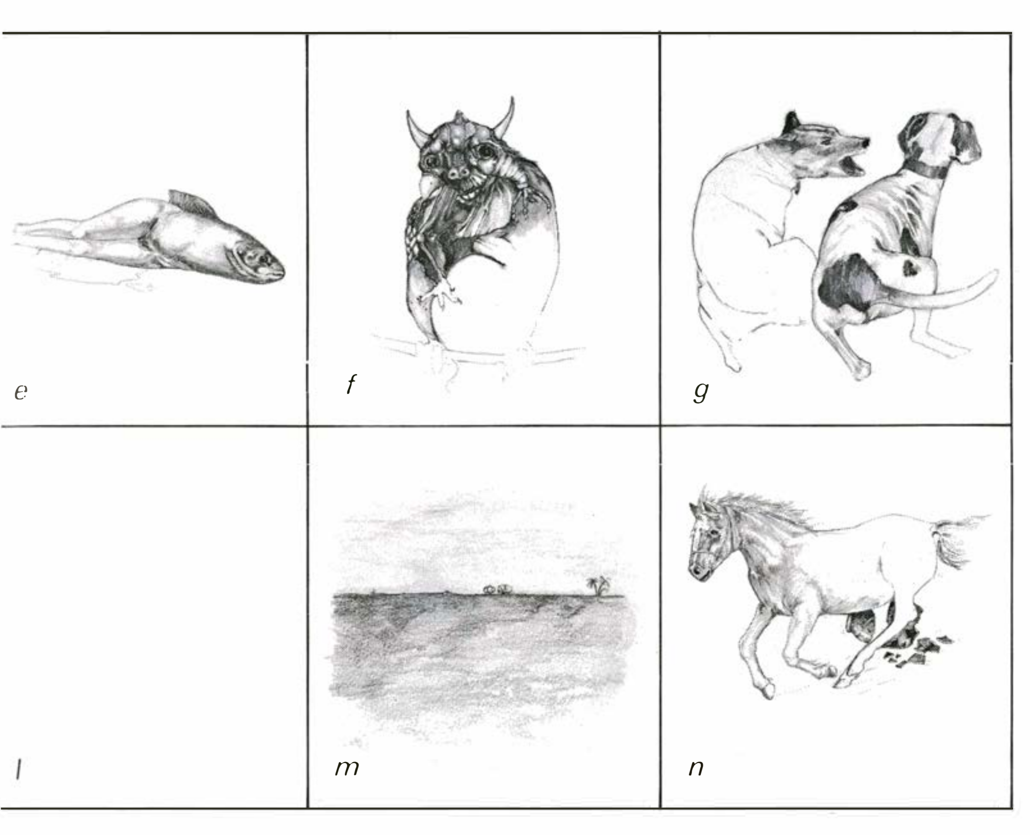
La relación R_H es reflexiva en A (pues cada elemento de A estará en la misma B_i que él mismo), simétrica (pues si x está en la misma subclase que y, y también estará en la misma que x) y transitiva. R_H es, pues, una relación de equivalencia.

En resumen, podemos decir que toda partición da lugar a una relación de equivalencia y que toda relación de equivalencia da lugar a una partición.

En química se usan diversas clasificaciones de los átomos. Consideremos la relación de equivalencia en que están dos átomos si y sólo si tienen exactamente el mismo número de protones y el mismo número de neutrones. Esta relación de equivalencia da lugar a la partición o clasificación del conjunto de los átomos en isótopos. Sin embargo, la relación de equivalencia en que están dos átomos si y sólo si tienen el mismo número de protones da lugar a la clasificación de los átomos en elementos químicos. Y, claro está, la clasificación de los átomos en elementos químicos determina unívocamente una relación de equivalencia en la que están dos átomos si y sólo si ambos pertenecen al mismo elemento. Esta relación coincide con la anterior, pues dos átomos pertenecen al mismo elemento químico si y sólo si poseen el mismo número de protones en su núcleo.

En la fonología de una lengua determinada podemos clasificar los sonidos emitidos y entendidos por los hablantes valiéndonos de la relación de equivalencia en que están dos sonidos si y sólo si son intercambiables sin que varíe el significado de la proferencia de que forman parte. Así, los sonidos de e abierta /ɛ/ y e cerrada /e/ son intercambiables en castellano –la proferencia /mɛ́sa/ significa lo mismo que /mésa/– pero no en francés –la proferencia /epɛ́/, espeso, significa algo distinto que /epé/, espada–. Esta relación de equivalencia da lugar a la clasificación de los sonidos de una lengua en fonemas.

En geometría euclídea, la relación de paralelismo es una relación de equivalencia entre las rectas del plano. Esta relación da lugar a la partición del conjunto de las rectas en direcciones. La dirección de una recta es precisamente la clase de equivalencia de esa recta respecto a la re-



esta clasificación, (*i*) que se agitan como locos, (*j*) innumerables, (*k*) dibujados con un pincel finísimo de pelo de camello, (*l*) etcétera, (*m*) que acaban de romper el jarrón, (*n*) que de lejos parecen moscas"

lación de paralelismo, es decir, la clase de todas las rectas paralelas a ella.

Vemos que la clasificación siempre tiene la misma estructura, aunque se establezca en ciencias tan distintas como la química, la fonología y la geometría. Cada átomo pertenece a uno y sólo un elemento. Cada sonido de una lengua pertenece a uno y sólo un fonema. Cada recta de un plano pertenece a una y sólo una dirección.

Clasificaciones: condiciones materiales de adecuación

En la práctica científica no sólo se exige que una clasificación satisfaga las condiciones formales de adecuación que acabamos de comentar, sino también que satisfaga ciertas condiciones materiales de adecuación peculiares de la ciencia de que se trate. Esto mismo suele expresarse en la pretensión de que la clasificación sea natural. Pero, ¿qué significa que una clasificación sea natural? Limitémonos a considerar el asunto en lo que atañe a la zoología. ¿Qué es una clasificación zoológica natural?

Podemos clasificar a los animales en tres clases: la de los que no llegan a los 2 años de vida, la de los que mueren entre los 2 y los 80 años y la de los que viven más de 80 años. Esto constituye una clasificación formalmente correcta de los animales. En efecto, los tres casos se dan, cada animal se encuentra en alguno de esos casos y ningún animal está a la vez en dos de esos casos. Sin embargo, esta clasificación sería rechazada por la comunidad de los zoólogos por no ser natural. ¿Por qué no es natural? ¿Y por qué es natural la clasificación de los animales en *phyla* (cordados, equinodermos, artrópodos, etc.)? La respuesta es que podemos enunciar muchas e interesantes leyes generales acerca de los artrópodos, por ejemplo, pero no acerca de los animales que viven entre 2 y 80 años. El identificar un animal concreto como artrópodo nos permite hacer muchas predicciones sobre ese animal, mientras que el identificarlo como viviendo entre 2 y 80 años no nos permite predecir gran cosa acerca de él.

En general, suele considerarse que una clasificación es más natural que otra si los conceptos que constituyen la primera son más fecundos teóricamente, en el sentido de que sirven para formular leyes más generales o más precisas o con más poder explicativo o predictivo. Esta es la razón que ya llevó a Aristóteles a incluir los cetáceos entre los mamíferos, y no entre los peces. Así resultaba posible formular leyes generales acerca de los peces –todos los peces son ovíparos, todos los peces son de sangre fría, todos los peces respiran por agallas, etcétera–, que no hubieran valido de haber incluido los cetáceos entre los peces.

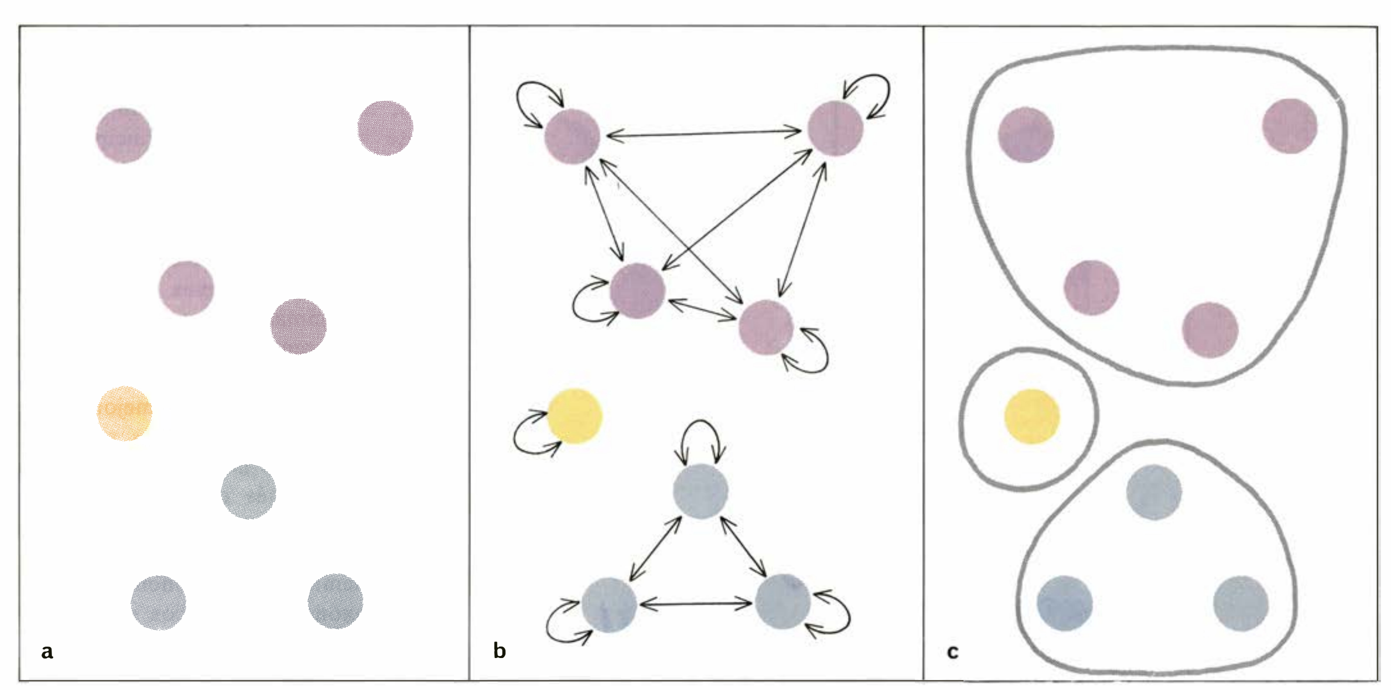
A la hora de concretar más lo que se entiende por clasificación natural en zoología, las opiniones discrepan. Actualmente hay dos escuelas principales de taxonomía: la evolutiva y la fenética. Según la taxonomía evolutiva, una clasificación natural ha de reflejar las relaciones filogenéticas entre los animales, agrupando en las mismas clases o taxones a los animales que están evolutivamente emparentados entre sí. Según la taxonomía fenética, la clasificación más natural será aquella que mejor refleje el parecido actual entre los animales, agrupando en los mismos taxones a los animales que más caracteres comunes compartan, con independencia de su genealogía.

La polémica entre taxónomos evolutivos y taxónomos fenéticos se extiende también a la cuestión de si el hombre inventa (como quieren los segundos) o más bien descubre (como pretenden los primeros) los diversos taxones biológicos y en especial las especies. Así escribe Ernst Mayr, uno de los más ilustres taxónomos evolutivos, que "los taxones inferiores no son colecciones arbitrarias, sino comunidades reproductivas mantenidas juntas por relaciones de cortejo y separadas de otras unidades similares no por las decisiones arbitrarias del clasificador, sino por los mecanismos aisladores codificados en el programa genético de cada organismo". Robert Sokal, uno de los iniciadores de la taxonomía fenética, propone por el contrario "basar las clasificaciones enteramente en el parecido, definiendo como clasificaciones naturales aquellas que determinan taxones cuyos miembros son en algún sentido más similares entre sí que con los miembros de los otros taxones. De este concepto de naturalidad se sigue que la clasificación natural será también la más predictiva".

No es éste el lugar de estudiar esta interesante polémica. Sólo nos importa señalar cómo, junto a las condiciones formales de adecuación de una clasificación, estructurales y comunes a todas las ciencias, en cada ciencia particular se suelen exigir condiciones materiales de adecuación o naturalidad, aunque, como muestra el caso de la biología, no siempre la comunidad científica está en completo acuerdo sobre en qué consista esa naturalidad.

Jerarquías de clasificaciones

Dadas dos clasificaciones del mismo dominio de objetos, a veces es posible compararlas en cuanto a finura y, a veces, no. Así, por ejemplo, la clasificación



PARTICION generada por una relación de equivalencia. En (*a*) vemos ocho objetos de diferentes colores. Las flechas de (*b*) representan la relación de equivalencia "... tiene igual color que..." entre objetos de (*a*). El espacio cociente de (*a*) por esa relación da lugar a una partición de los objetos de (*a*) en clases de equivalencia respecto a esa relación, clases de equivalencia que aparecen delimitadas en (*c*).

de los primates en prosimios y simios no es comparable con la clasificación de los mismos en machos y hembras. La clasificación de los libros por su fecha de publicación no es comparable con su clasificación por el lugar de su impresión. Sin embargo, la clasificación de los mamíferos en familias sí es comparable con su clasificación en órdenes. La primera es más fina que la segunda. Y la clasificación del territorio nacional por municipios es más fina que su clasificación en provincias.

Sean $A=\{A_1,...A_n\}$ y $B=\{B_1,... B_m\}$ clasificaciones o particiones del mismo dominio D. Entonces podemos decir que A es tanto o más fina que B si y sólo si para cada $A_i \epsilon A$ y cada $B_j \epsilon B$ ocurre que $A_i \subset B_j$ o $A_i \cap B_j = \emptyset$.

Suele ser característico de las ciencias en que los conceptos clasificatorios desempeñan un papel importante el que las clasificaciones no aparezcan solas, sino que se usen diversas clasificaciones de finura decreciente del mismo dominio, engarzadas entre sí y formando jerarquías, donde por jerarquía entendemos una sucesión de clasificaciones comparables entre sí y de finura decreciente.

Más precisamente, decimos que H es una jerarquía taxonómica sobre D si y sólo si hay B_1 ... B_n, tales que: (1) $H = \{B_1 ... B_n\}$; (2) para cada i ($1 \leqslant i \leqslant n$): B_i es una partición de D; (3) para cada i ($1 \leqslant i \leqslant n$): B_i es tanto o más fina que B_{i+1}.

Aquí D es el dominio básico de individuos, y cada B_i es una categoría de nivel i. En general, una jerarquía sobre D es una clase de categorías sobre D. Una categoría sobre D es una partición de D, es decir, una clase de taxones de D. Y cada taxón de D es una clase de elementos de D.

La jerarquía taxonómica más conocida es la jerarquía procedente de Linné para la clasificación de los seres vivos. La jerarquía linneana L abarca 7 categorías, cada una de las cuales es una partición del conjunto de los seres vivos. L={especie, género, familia, orden, clase, phylum, reino}.

Cada ser vivo es miembro de un taxón de cada una de esas siete categorías. Así, el perro Lassie es a la vez miembro del taxón *Canis familiaris* (de la categoría especie), del taxón *Canis* (de la categoría género), del taxón *Canidae* (de la categoría familia), del taxón *Carnivorae* (de la categoría orden), del taxón *Mamalia* (de la categoría clase), del taxón *Chordata* (de la categoría phylum) y del taxón *Animalia* (de la categoría reino).

Aunque las jerarquías taxonómicas más conocidas son las de la biología, también en otras ciencias nos encontramos con diversas clasificaciones de finura decreciente que forman jerarquías. Pensemos en las clasificaciones de los átomos en química. Una clasificación muy fina es la partición de los átomos en isótopos, otra menos fina es su partición en elementos, otra aún menos fina es su partición en grupos o familias. Por tanto, H={isótopo, elemento, grupo} constituye una jerarquía sobre el dominio de todos los átomos. Cada átomo es miembro de un taxón de cada una de esas categorías. Así, por ejemplo, un átomo determinado puede ser a la vez miembro del taxón *potasio-39* (de la categoría isótopo), del taxón *potasio* (de la categoría elemento) y del taxón *alcalino* (de la categoría grupo).

Conceptos comparativos

Así como los conceptos clasificatorios corresponden a los sustantivos y adjetivos del lenguaje ordinario, así también los conceptos comparativos encuentran su punto de partida en un rasgo de nuestra lengua cotidiana: el llamado por los gramáticos grado comparativo de los adjetivos. El lenguaje ordinario no sólo nos permite "clasificar" a nuestros congéneres en altos y bajos; también nos permite precisar que un individuo determinado, aunque bajo, es más alto que otro. Expresiones como "más alto", "más viejo", "mayor", "más ligero", etcétera, corresponden a conceptos comparativos.

Introducir un concepto comparativo para una característica que los individuos de un dominio poseen en mayor o menor grado consiste en definir dos relaciones de coincidencia y precedencia respecto a esa característica, es decir, indicar cuándo dos objetos de ese dominio coinciden respecto a esa característica y cuándo uno precede al otro respecto a ella. Un concepto comparativo sirve así para establecer comparaciones en más y menos. Si identificásemos los conceptos cualitativos con los clasificatorios y los cuantitativos con los métricos, resultaría que en la ciencia se usan otros tipos de conceptos además de los cualitativos y cuantitativos: los conceptos comparativos (o topológicos). Los conceptos comparativos no sólo permiten diferenciar

más finamente que los clasificatorios, sino que además representan un primer paso para la posterior introducción de conceptos métricos.

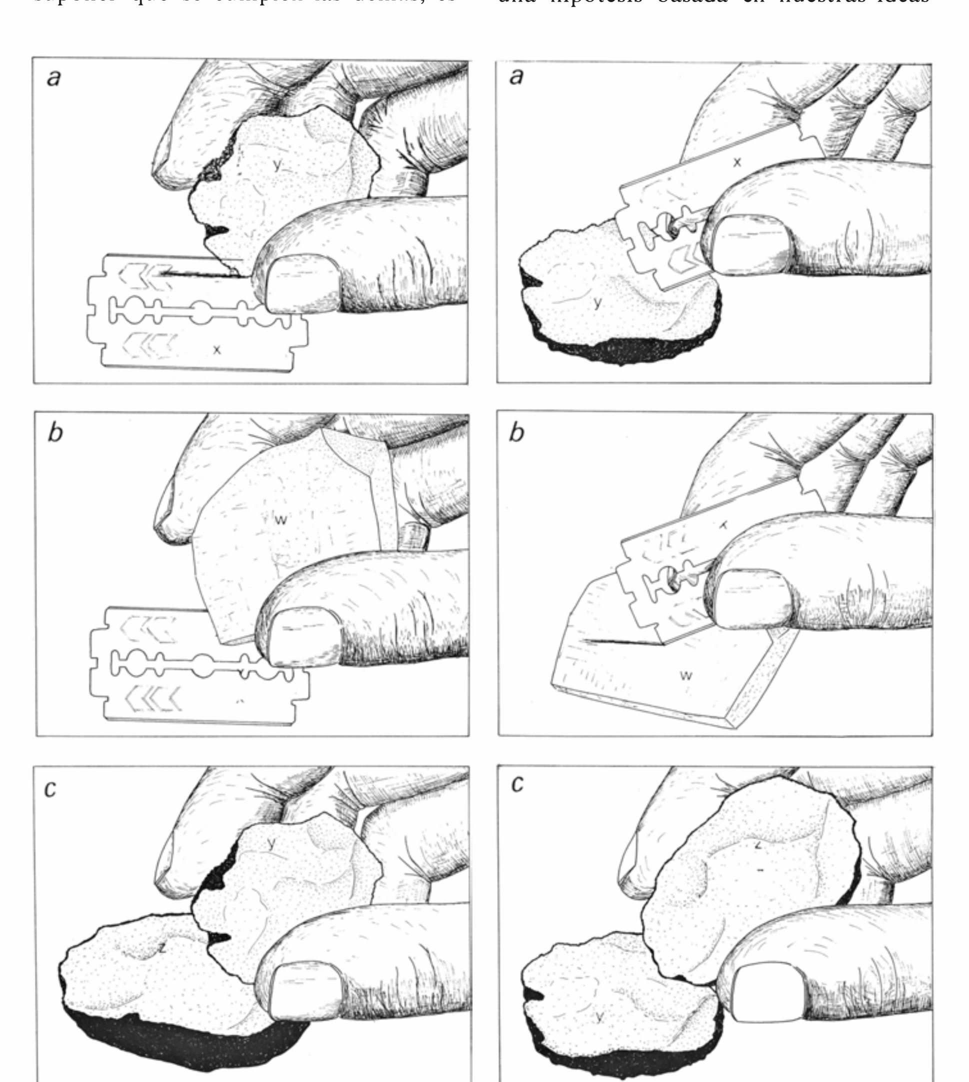
Llamemos K y P a las relaciones de coincidencia y precedencia respecto a una característica determinada que los objetos de un dominio A poseen en mayor o menor grado. El concepto comparativo < K, P> ha de cumplir ciertas condiciones formales de adecuación para ser científicamente aceptable. En primer lugar, K ha de ser una relación de equivalencia en A (es decir, todo objeto ha de coincidir consigo mismo respecto a la característica de que se trate; si un objeto coincide con otro, entonces también el otro con el uno; y si uno coincide con otro y ese otro con un tercero, entonces el primero ha de coincidir con el tercero). P ha de ser transitiva en A (es decir, si un objeto es más –respecto a la característica en cuestión– que otro y ese otro más que un tercero, entonces el primero es más que el tercero). Además P ha de ser K-irreflexiva (es decir, el que un objeto coincida con otro respecto a la característica estudiada excluye que sea mayor o menor que él respecto a esa misma característica). Finalmente, todos los miembros de A han de ser comparables respecto a < K, P> (es decir, dados dos objetos cualesquiera, o bien coinciden entre sí, o bien uno de ellos es más o menos que el otro respecto a la característica de que se trate).

Podemos resumir las condiciones formales de adecuación de un concepto comparativo < K, P> en un dominio A exigiendo que < A, K, P> constituya un sistema comparativo. < A, K, P> es un *sistema comparativo* si y sólo si para cualesquiera elementos x, y, z de A ocurre que: (1) $x\mathrm{K}x$; (2) si $x\mathrm{K}u$, entonces $u\mathrm{K}x$; (3) si $x\mathrm{K}u$ y $u\mathrm{K}z$, entonces $x\mathrm{K}z$; (4) si $x\mathrm{P}u$ y $u\mathrm{P}z$, entonces $x\mathrm{P}z$; (5) si $x\mathrm{K}z$, entonces no $x\mathrm{P}z$, y (6) $x\mathrm{K}z$ o $x\mathrm{P}z$ o $z\mathrm{P}x$.

Un ejemplo típico de concepto comparativo es el concepto de *dureza* usado en mineralogía. Este concepto comparativo de dureza sobre el dominio de los minerales se basa en la prueba del rayado. Dados dos minerales, x y z, decimos que x es más duro que z si y sólo si x raya a z, pero z no raya a x. Y decimos que x coincide respecto a dureza con z si ocurre que ni x raya a z ni z raya a x (o x y z son el mismo mineral). Este concepto comparativo de dureza cumple las 6 condiciones formales de adecuación formuladas en la definición de sistema comparativo. Las condiciones (1) y (5) se cumplen por definición. El que las otras condiciones también se cumplan constituye una hipótesis empírica (por ahora bien confirmada) de la mineralogía.

Otro ejemplo de concepto comparativo (uno que sirve de primer paso para la posterior introducción de un concepto métrico) es el concepto (premétrico) de *masa*. Este concepto comparativo de masa tiene como dominio el ámbito de los cuerpos manejables (es decir, ni demasiado pequeños ni demasiado grandes, sino manipulables con la mano) y se basa en la prueba de la balanza. Dados dos objetos x y z, decimos que x coincide respecto a masa con z, si, colocados ambos en sendos platillos de una balanza, ésta permanece equilibrada (o bien si x y z son el mismo cuerpo). Y decimos que x tiene más masa que z, si colocados ambos en sendos platillos de la balanza, ésta se desequilibra a favor del platillo donde hemos colocado x. Este concepto comparativo de masa cumple las 6 condiciones formales de adecuación. Y también en este caso ocurre que (1) y (5) se cumplen por definición, mientras que, al suponer que se cumplen las demás, estamos haciendo diversas hipótesis (por lo demás triviales, claro) tanto sobre el comportamiento de la naturaleza como sobre el buen estado de nuestra balanza.

En paleontología se emplea un concepto comparativo de *antigüedad* cuando ocurre que resulta difícil datar absolutamente los fósiles hallados en un yacimiento estratificado. El dominio de ese concepto comparativo de antigüedad está constituido por los fósiles que se encuentran en los diversos estratos geológicos del yacimiento. Decimos que un fósil x coincide respecto a antigüedad con un fósil z si y sólo si x y z se encuentran en el mismo estrato. Y decimos que x es más antiguo que z si x se encuentra en un estrato inferior a aquel en el que se encuentra z. También este concepto cumple las 6 condiciones formales de adecuación, y como en los otros casos, también aquí (1) y (5) se cumplen por definición, y que se cumple el resto es una hipótesis basada en nuestras ideas



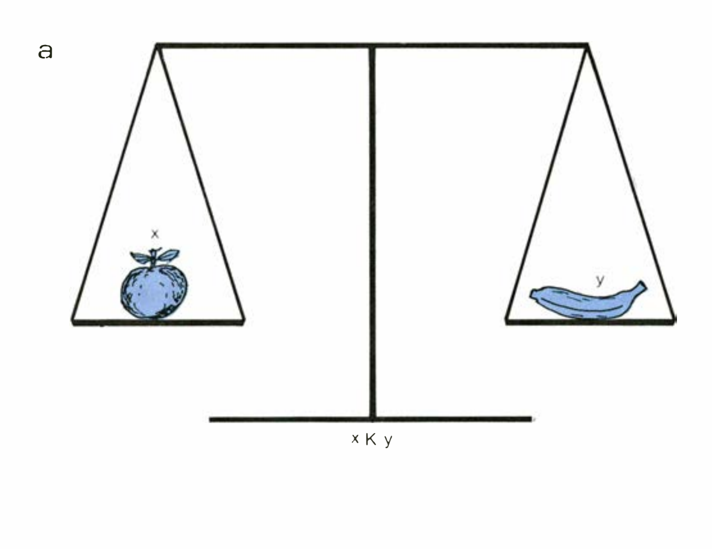
EL CONCEPTO COMPARATIVO DE DUREZA puede ser establecido mediante la prueba del rayado. En (*a*) vemos que *y* (un trozo de sílex) raya a *x* (una hoja de afeitar de acero), pero no a la inversa: *x* es más blando que *y*. En (*b*) es *x* el que raya a *w* (un trazo de calcita), sin ser rayado por él: *w* es más blando que *x*. En (*c*) aparecen dos trozos distintos de sílex *y* y *z*, ninguno de los cuales raya al otro ni es rayado por el otro: *y* coincide con *z* en cuanto a dureza.

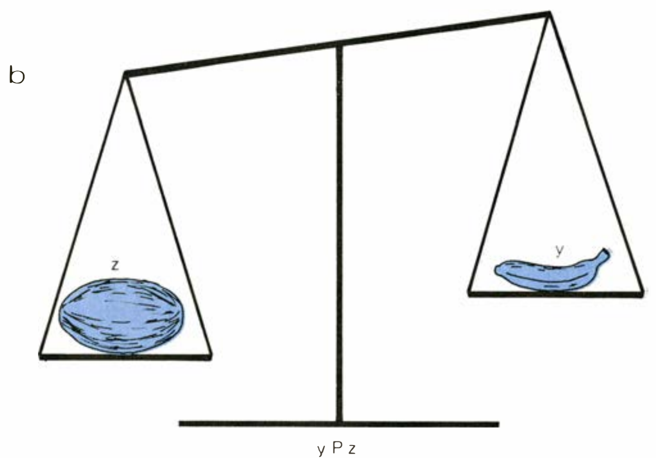
acerca de la formación de las rocas sedimentarias y la fosilización de los restos de seres vivos.

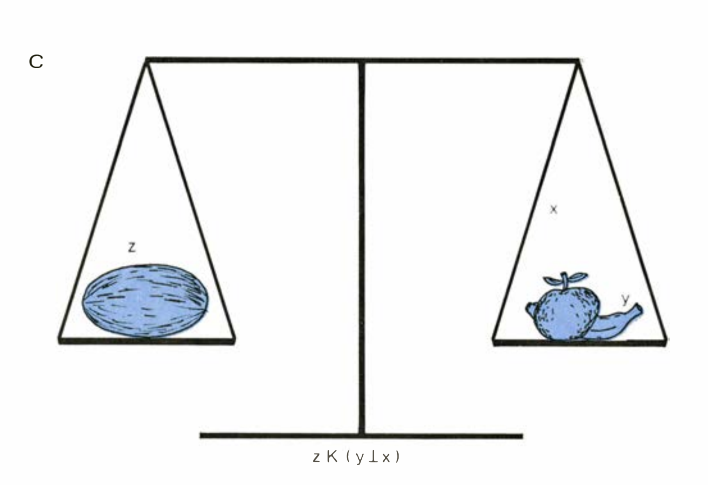
Cuando queremos precisar más nuestras nociones acerca de un ámbito determinado, con frecuencia resulta más fácil introducir un concepto comparativo que otro métrico. Así podríamos tratar de precisar el concepto de fortaleza (muscular) en un dominio de seres humanos mediante la prueba de echarse un pulso (sería más fuerte que otro el que, echándose un pulso, derribase al otro; coincidirían los que, echándose un pulso, ni derribasen ni fuesen derribados). Pero estaría por ver si este concepto cumple más o menos las condiciones formales de adecuación, y si sirve para algo.

El concepto de metal es en principio clasificatorio. Clasificamos los elementos químicos en metales y no-metales. Pero al definir lo que entendemos por metal (elemento que posee en la capa más externa de la corteza un número pequeño de electrones, de los que puede desprenderse fácilmente, dando lugar a iones positivos; presenta gran conductividad eléctrica y calórica, etc.), es evidente que unos elementos poseen esas características en un grado mayor que otros. Algunos elementos (como los alcalinos) son "muy metales"; otros (como los halógenos) no son nada metales; los demás ocupan grados intermedios. El mismo estaño se comporta en una de sus formas como metal, y, en otra, como no metal. Por ello, podríamos tratar de reformular nuestra noción de *metalidad* como concepto comparativo, explicitando criterios que nos sirvan para decidir, de dos elementos cualesquiera, si coinciden respecto a metalidad o uno es más metálico que el otro. Y hemos visto cómo, además del concepto métrico de masa, hay un concepto (previo) comparativo de masa. El punto a retener es que el ser clasificatorio, comparativo o métrico, como el ser cualitativo o cuantitativo, no son propiedades de las cosas, sino de los conceptos que empleamos para pensar en las cosas y hablar de las cosas.

Señalemos, finalmente, que si bien no siempre es fácil (ni posible) pasar de un sistema clasificatorio a otro comparativo, la inversa (es decir, pasar de un concepto comparativo a una clasificación) siempre es posible, fácil e incluso trivial. En efecto, sea $<K, P>$ un concepto comparativo sobre un dominio D. La relación de coincidencia K (que, como sabemos, es una relación de equivalencia) determina entonces unívocamente una partición o clasificación de D, a saber, D/K, es decir, el conjunto cociente de D respecto a K. Esta partición, además de soler ser de gran finura, tiene la ventaja de estar (irreflexivamente) ordenada por la relación P, con lo que obtenemos una mayor información sobre las interrelaciones mutuas entre las diversas clases que constituyen esa clasificación.







EL CONCEPTO COMPARATIVO DE MASA puede ser establecido mediante la prueba de la balanza. En (*a*) colocamos dos objetos *x* (una manzana) e *y* (un plátano) en los platillos y la balanza permanece equilibrada: *x* coincide con *y* en cuanto a masa. En (*b*) colocamos *z* (un melón) e *y* en los platillos y la balanza se inclina hacia *z*: *y* precede a *z* en cuanto a masa, es decir, *y* es más ligero que *z*. En (*c*) *z* coincide en cuanto a masa con la combinación de *y* y *x*.

Conceptos métricos

Los conceptos métricos, también llamados conceptos cuantitativos o magnitudes, no tienen correspondencia en el

lenguaje ordinario. Son una creación original de los lenguajes científicos. Son característicos de los estadios más avanzados de la ciencia. Piénsese que la revolución científica del siglo XVII consistió en gran parte en la introducción y el uso sistemático de los conceptos métricos en la física, que durante los dos mil años anteriores había estado basada en los conceptos cualitativos.

Los conceptos métricos asignan números reales o vectores a objetos o sucesos. Los conceptos métricos –como *masa* o *tiempo*– que asignan números reales a determinados objetos o sucesos se llaman *magnitudes escalares.* Los conceptos métricos –como *fuerza* o *velocidad*– que asignan vectores se llaman *magnitudes vectoriales.* Para simplificar nuestro tratamiento, vamos a limitarnos aquí a hablar de las magnitudes escalares, aunque *mutatis mutandis* lo mismo podría ser dicho de las vectoriales. Cuando en lo sucesivo hablemos de concepto métrico queremos decir concepto métrico escalar.

En una primera aproximación podemos decir que un concepto métrico f en un dominio A es simplemente f: A→**R**, es decir, una aplicación del dominio A sobre el conjunto de los números reales o, con otras palabras, una asignación de un número real a cada uno de los objetos de A. Así, el concepto métrico de *masa* asigna un número real a cada cuerpo, el de *longitud* asigna un número real a cada dos señales en una superficie plana de un cuerpo o a cada dos cuerpos, el de *tiempo* asigna un número real a cada dos sucesos, el de *frecuencia* asigna un número real a cada onda, el de *resistencia* asigna un número real a cada conductor eléctrico, el de *índice cefálico* asigna un número real a cada cabeza, el de *producto nacional bruto* asigna un número real a cada economía nacional y año, el de *tasa de natalidad* asigna un número real a cada población y año, y así sucesivamente.

En una segunda aproximación podemos observar que con frecuencia tratamos de introducir un concepto métrico en un ámbito en el que ya disponemos de un concepto comparativo. La *metrización* de un ámbito o de una característica consiste precisamente en la introducción de un concepto métrico en ese ámbito o para esa característica. (No hay que confundir metrización y medida. La *medida* supone que ya disponemos de un concepto métrico y consiste en la búsqueda del número real o vector que ese concepto métrico asigna a un objeto o suceso determinado.) Muchas veces, de lo que se trata es de metrizar un ámbito ya previamente ordenado, es decir, se trata de metrizar un sistema comparativo o, dicho todavía con otras palabras, se trata de introducir un concepto métrico para algo para lo que ya disponemos de un concepto comparativo. Si $< A, K, P>$ es un sistema comparativo que pretendemos metrizar mediante la función f:A→**R**, lo primero que debemos exigir es que f conserve el orden establecido por $< K, P>$ en A, es decir, que f asigne el mismo número real a los objetos coincidentes y que, si un objeto precede a otro, entonces f asigne un número real menor al primer objeto que al segundo. Más precisamente, la condición formal de adecuación de un concepto métrico f que pretenda metrizar el sistema $< A, K, P>$ de la que estamos hablando exige que para cada dos objetos x y z de A ocurra que (1) si xKz, entonces $f(x)=f(z)$, y (2) si xPz, entonces $f(x)< f(z)$.

Un concepto métrico de este tipo no sólo asignará números a las cosas, sino que además nos ofrecerá cierta información sobre el orden en que están esas cosas respecto a la característica que hayamos metrizado. Aquí lo que habremos hecho habrá sido representar determinadas características cualitativas o empíricas de los objetos del dominio A (de personas, minerales, poblaciones o lo que sea) por características cuantitativas o matemáticas de los números reales. Lo que habremos hecho será, pues, establecer un homomorfismo entre el sistema empírico comparativo $< A, K,P>$ y el sistema numérico $< \mathbf{R}, =, < >$, donde **R** es el conjunto de los números reales y $=$, $<$ son la identidad y la relación "menor que" entre números reales.

Esta representación de un sistema empírico en otro numérico constituye la esencia del concepto métrico. En una tercera aproximación, podemos decir que un concepto métrico f es un homomorfismo de un sistema empírico en un sistema numérico homólogo. Un *sistema* está constituido por un dominio de individuos y una serie de relaciones y funciones en ese dominio. Dos sistemas son *homólogos* si tienen el mismo número de relaciones y de funciones y si los números arios se corresponden (es decir, si la primera relación de un sistema es binaria, también lo es la del otro, etcétera).

Sean $\mathscr{A} = < A, R_1...R_n, g_1...g_m >$ y $\mathscr{R} = < \mathbf{R}, S_1...S_n, h_1...h_m >$ dos sistemas homólogos. Decimos que f es un *homomorfismo* de $\mathscr{A}$ en $\mathscr{R}$ si y sólo si ocurren las siguientes tres cosas: (1) f es una función que a cada objeto de A asigna un número real, es decir, f: A→**R**. (2) Siempre que los objetos a ...a_n de A están entre sí en la relación R_i, los correspondientes números reales $f(a)$... $f(a_n)$ están también entre sí en la correspondiente relación S_i. (3) Siempre que la función g_i de $\mathscr{A}$ asigna a cada n objetos $a_1...a_n$ de A otro objeto a_{n+1} de A, la correspondiente función h_i de $\mathscr{R}$ asigna a los correspondientes números reales $f(a_r)..f(a_n)$ el correspondiente número real $f(a_{n+1})$.

Con esto queda precisado lo que entendemos por concepto métrico: un concepto métrico es un homomorfismo entre un sistema empírico y un sistema numérico. El análisis estructural de la metrización de un sistema empírico suele constar de cuatro pasos: (1) Definición del sistema empírico; (2) formulación de axiomas o hipótesis que expresan ciertas características cualitativas de ese sistema empírico; (3) prueba de un teorema de representación, que afirma la existencia de un homomorfismo de ese sistema empírico en cierto sistema numérico; (4) prueba de un teorema de unicidad, que indica hasta qué punto el homomorfismo es unívoco, es decir, cuáles son las transformaciones (llamadas permisibles) del homomorfismo dado, que también constituyen homomorfismos del mismo sistema empírico en el mismo sistema numérico.

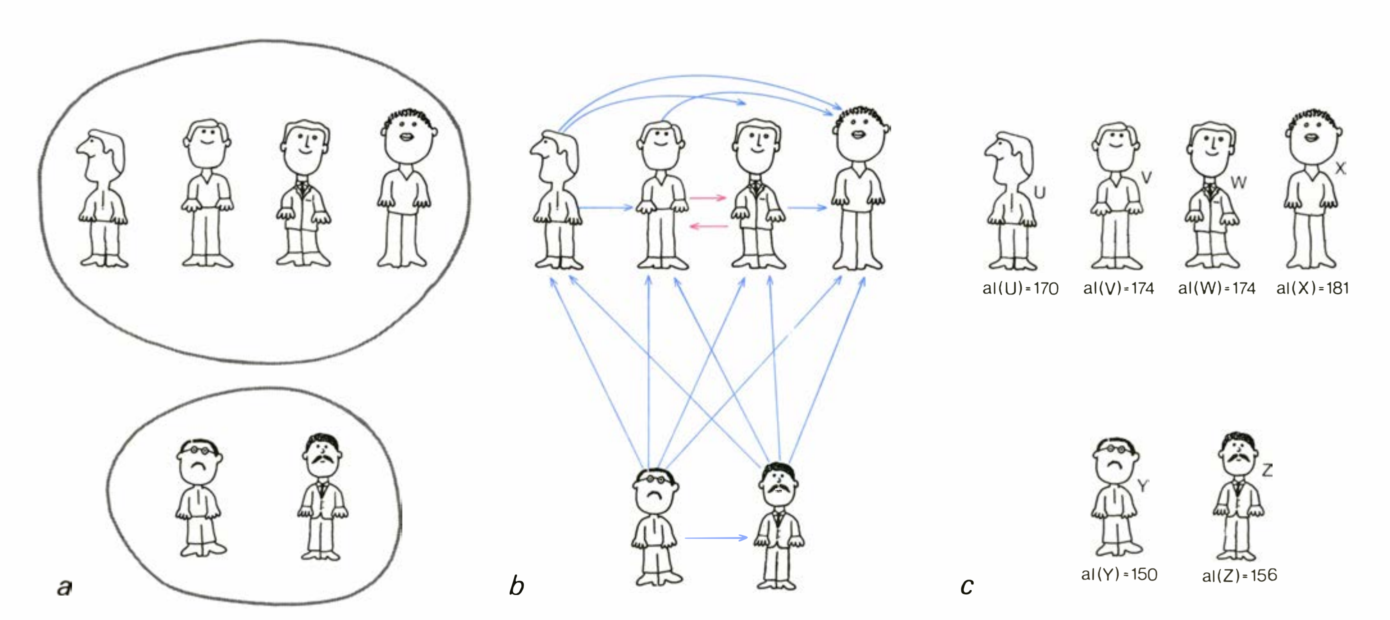
Aquí no vamos a desarrollar este análisis, sino que nos limitaremos a distinguir y ejemplificar algunos de los principales tipos de conceptos métricos.

A veces se identifica el concepto métrico con una escala, pero quizá sea más conveniente identificar una *escala* con un homomorfismo concreto de un sistema empírico en un sistema numérico, y el *concepto métrico* con la clase de todos los homomorfismos del primer sistema en el segundo. Así, para un concepto métrico dado varias transformaciones de escalas serían permisibles. Y el hecho de que un mismo concepto métrico pueda expresarse en varias escalas corresponde evidentemente a la práctica científica.

Escalas ordinales

Las escalas ordinales son las más pobres desde el punto de vista de la información que nos suministran. De hecho, su rendimiento teórico no es mayor que el de los conceptos comparativos. Se limitan a asignar números, conservando el orden de un sistema comparativo dado.

La escala de Richter para la intensidad de los terremotos, la de Beaufort para la de los vientos y la de Mohs para la dureza de los minerales son típicos ejemplos



SER CLASIFICATORIO, COMPARATIVO O METRICO es una propiedad de los conceptos empleados para hablar de las cosas, y no de las cosas mismas. Así, de la misma característica objetiva (la altura de diversos individuos) podemos hablar tanto mediante los conceptos clasificatorios "alto" y "bajo" representados en (*a*), como mediante el concepto comparativo "... es más bajo que..." (*flecha azul*) y "... es tan alto como..." (*flecha roja*), representado en (*b*), como mediante un concepto métrico de altura simbolizado por el factor *al*, indicado en (*c*).

de escalas ordinales. Consideremos la última de las citadas.

Como ya vimos anteriormente, en mineralogía se dispone de un concepto comparativo de dureza basado en la prueba del rayado. Siempre que asignemos números a los minerales de tal manera que a dos minerales les asignemos el mismo número o a uno de ellos un número menor que al otro según que coincidan en cuanto a dureza o el uno sea menos duro que el otro conforme a la prueba del rayado, tendremos una escala ordinal de dureza. Así, el minerólogo alemán Friedrich Mohs, en 1822, decidió asignar números a algunos minerales, estableciendo así la "escala de Mohs". En concreto asignó el 1 al talco, el 2 al yeso, el 3 a la calcita, el 4 a la fluorita, el 5 al apatito, el 6 a la ortosa, el 7 al cuarzo, el 8 al topacio, el 9 al corindón y el 10 al diamante. Si un mineral por la prueba del rayado resulta ser, por ejemplo, más duro que el cuarzo y más blando que el topacio, se le asigna un número intermedio entre el 7 y 8, como el 7,5.

Sea M el conjunto de los minerales. Sean K y P las relaciones de coincidencia respecto a dureza y de menor dureza según la prueba del rayado. La escala de Mohs es un homomorfismo f del sistema empírico $< M, K, P >$ en el sistema numérico $< \mathbf{R}, =, < >$, tal que f(talco)= =1, f(yeso)=2, f(calcita)=3, f(fluorita)=4, y así sucesivamente.

La escala de Mohs se limita a expresar numéricamente el hecho de que un mineral es más, o menos, duro que otro, pero no nos dice *cuánto* más, o menos, duro es que el otro. No mide diferencias de dureza. Esta limitación es común a todas las escalas ordinales. Precisamente por ello, son muchas las transformaciones permisibles, es decir, las transformaciones del homomorfismo que dan lugar a homomorfismos del mismo tipo.

Sean f y h dos funciones que asignan números reales a los elementos de un dominio A. Decimos que h es una *transformación monótona* de f si para cada dos elementos x y z de A ocurre que si $f(x)< f(z)$, entonces $h(x)< h(z)$. Pues bien, si f es un homomorfismo de un sistema empírico en un sistema numérico y constituye una escala ordinal, cualquier transformación monótona de f será también un homomorfismo del mismo sistema empírico en el mismo sistema numérico y, por tanto, será igualmente una escala ordinal. Si en vez de asignar 1 al talco, 2 al yeso, 3 a la calcita, 4 a la fluorita, etc., como hacía Mohs, asignamos 0 al talco, 500 al yeso, 500,5 a la calcita, 507 a la fluorita, etc., esa asignación sigue siendo un homomorfismo de $< M, K, P >$ en $< \mathbf{R}, =, < >$, es decir, sigue siendo una escala ordinal de dureza. Precisamente esta indeterminación es la que impide que pueda haber una fórmula general para pasar de una escala ordinal a otra (correspondiente al mismo concepto).

Escalas proporcionales

Las escalas proporcionales son las más ricas desde el punto de vista de la información que suministran. No sólo nos dicen que un objeto es más, o menos, que otro respecto a alguna característica, sino que nos señalan en qué proporción exacta el uno es más, o menos, eso que el otro.

Las escalas correspondientes a los conceptos básicos de la física, como *masa*, *longitud* o *tiempo*, son escalas proporcionales, que, además, constituyen magnitudes aditivas o extensivas, por disponer en sus respectivos sistemas empíricos de una operación correspondiente a la adición.

Ya habíamos aludido al concepto comparativo de masa, basado en la prueba de la balanza y aplicable al dominio de los objetos físicos manejables. Consideremos ahora la operación empírica consistente en colocar dos objetos juntos (que por convención dan lugar a un nuevo objeto) en el mismo platillo de la balanza, y designemos esta operación mediante el signo $\perp$. Sea $< A, K, P, \perp >$ el sistema empírico formado por el conjunto de los objetos físicos manejables, las relaciones de coincidencia y precedencia respecto a la prueba de la balanza y la operación de colocar juntos dos objetos en el mismo platillo, de la que acabamos de hablar. Una escala de masa es un homomorfismo de $< A, K, P, \perp >$ en $< \mathbf{R}, =, <, + >$, es decir, una función f:A→**R**, tal que para cualesquiera x y z de A: (1) si xKz, entonces $f(x)$= =$f(z)$; (2) si xPz, entonces $f(x)< f(z)$ y (3) $f(x \perp z)=f(x) + f(z)$.

Hay muchas funciones que cumplen esas condiciones, muchas escalas. ¿Cómo fijar una? Eligiendo un objeto cualquiera

de A y asignándole convencionalmente un número cualquiera. Así, en la escala métrica decimal se elige un determinado cilindro de platino e iridio (el "kilo patrón") que se conserva en el museo de pesas y medidas de Sèvres y se le asigna el número 1000. Con esto queda fijada la escala de masa en gramos.

A diferencia de lo que pasaba con las escalas ordinales, no todas las transformaciones monótonas de escalas proporcionales dan lugar a escalas proporcionales. Supongamos que un frasco destapado tiene 200 gramos de masa, y su tapa, 100 gramos. Por tanto, el frasco tapado tendrá 300 gramos de masa. Una transformación monótona h de la escala métrica decimal en gramos m podría asignar al frasco el número 2, a su tapa, el 1, y al frasco tapado, el 9. Pero esa función no sería un homomorfismo de $< A, K, P, \perp >$ en $< \mathbf{R}, =, <, + >$. En efecto, mientras que m (frasco $\perp$ tapa)$=m$ (frasco)$+m$ (tapa)$=200+100=300$, ocurriría que h (frasco $\perp$ tapa)$=9 \neq h$(frasco)$+h$(tapa)$=$ $=2+1=3$.

En realidad, la mayoría de las transformaciones monótonas de una escala proporcional *no* son escalas proporcionales. Sólo las transformaciones similares dan lugar de nuevo a escalas proporcionales.

Sea f una función que asigna números reales a los elementos de A. Una función h: A$\rightarrow$**R** es una *transformación similar* de f si y sólo si hay un número positivo fijo k tal que para cada objeto x de A ocurre que $h(x)=k \cdot f(x)$, es decir, $h(x)$ es siempre el producto de $f(x)$ por un número positivo fijo. Pues bien, un homomorfismo f de un sistema empírico en un sistema numérico constituye una *escala proporcional* si y sólo si cualquier transformación similar de f es también un homomorfismo del mismo sistema empírico en el mismo sistema numérico.

De aquí se sigue que para pasar de una escala proporcional a otra basta siempre con multiplicar por un número fijo. Así, para pasar de una escala en kilos a otra en gramos basta con multiplicar por 1000; para pasar de una escala en libras a otra en kilos basta con multiplicar por 0,453, etcétera.

Para citar otro ejemplo de escalas proporcionales, consideremos el concepto métrico de longitud. Sea B el conjunto de las barras metálicas. Sea ahora Δ la operación de concatenación de barras a lo largo de una línea recta, es decir, la operación de colocar una barra a continuación de otra. Ahora podemos definir una escala de longitud como un homomorfismo del sistema empírico $< B, \Delta >$ en el sistema numérico $< \mathbf{R}, + >$. Como hay muchas funciones f:B$\rightarrow$**R** que constituyen homomorfismos de $< B, \Delta >$ en $< \mathbf{R}, + >$, para fijar una hemos de elegir un objeto de B y asignarle un número determinado. Así, en la escala métrica decimal se eligió (en 1889) una barra de platino-iridio conservada en el museo de pesas y medidas de Sèvres y se le asignó el número 1.

Como en el caso de la masa, y puesto que aquí también nos las habemos con escalas proporcionales, con la longitud ocurre que no toda transformación monótona, sino sólo toda transformación similar conduce de una escala de longitud a otra. Por ello, para pasar de una escala de longitud a otra basta con multiplicar por un número fijo. Así, para pasar de una escala en millas a otra en metros, basta con multiplicar por 1609; para pasar de una escala en metros a otra en centímetros, basta con multiplicar por 100, y así sucesivamente.

Téngase en cuenta que aquí hemos introducido la longitud sólo respecto a barras metálicas (o la masa respecto a objetos manejables). Esto no es sino el primer paso para luego ir extendiendo estos conceptos mediante leyes científicas a ámbitos más amplios, o, si se prefiere, éstos no son sino los primeros de una sucesión de conceptos métricos de masa y longitud de alcance creciente. Téngase también en cuenta que nos hemos limitado a señalar las condiciones formales de adecuación de los conceptos métricos (resumidas en la exigencia de que constituyan homomorfismos de sistemas empíricos en sistemas numéricos), dejando de lado las condiciones materiales de adecuación que, por ejemplo, en la física han llevado a una constante revisión de los objetos patrones o estándar que sirven para fijar las escalas. Así, y puesto que acabamos de hablar del concepto de longitud, podemos recordar cómo la comunidad de los físicos ha ido pasando de una escala de longitud basada en la diezmillonésima parte del cuadrante del meridiano terrestre que pasa por Dunkerke y Barcelona (1799), a otra basada en la barra del museo de Sèvres (1899) y a otra (en nuestros días) basada en la longitud de onda de la radiación emitida por el isótopo krypton-86, excitado a la temperatura del triple punto del nitrógeno (–210° C.).

Magnitudes extensivas e intensivas

Hemos visto que los conceptos de masa o de longitud son (clases de) homomorfismos de un sistema empírico que contiene una operación binaria de combinación de objetos (la colocación de dos objetos juntos en la balanza, la concatenación de barras una a continuación de otra) en un sistema numérico que contiene la adición. Las magnitudes de este tipo se llaman *magnitudes aditivas* o *extensivas*. Lo esencial de una magnitud aditiva f estriba en la correspondencia entre la operación binaria de combinación y la adición. Si a la primera la designamos por $\perp$, siempre ocurre que para cada dos individuos x y z del dominio:

$$f(x \perp y)=f(x)+f(y)$$

Así, la masa de un objeto compuesto de dos partes es igual a la suma de las masas de sus partes. La longitud del objeto resultante de colocar dos objetos en línea recta uno a continuación de otro es igual a la suma de sus longitudes. Esto no sólo ocurre con la masa o la longitud. Lo mismo ocurre con el tiempo (si un proceso se divide en dos partes tales que la segunda se inicia al acabarse la primera, la duración del proceso global es igual a la suma de las duraciones de sus partes). El tiempo es también una magnitud aditiva o extensiva.

Es necesario explicitar exactamente las operaciones del sistema empírico para poder determinar si un concepto métrico que lo represente sobre un sistema numérico es una magnitud aditiva o no. Consideremos el caso de la resistencia eléctrica. Como es bien sabido, en un circuito podremos colocar varias "resistencias", es decir, varios conductores, en serie o en paralelo. Sea C el conjunto de los conductores eléctricos, sea M la relación en que está un conductor con otro cuando el primero ofrece menor o igual resistencia a la corriente eléctrica que el segundo y sea $\perp$ la operación de colocar conductores en serie. El concepto métrico de resistencia es una magnitud aditiva, pues sus escalas son homomorfismos de $< C, M, \perp >$ en $< \mathbf{R}, \leqslant, + >$. Si en vez de considerar la operación de colocar conductores en serie, hubiéramos elegido la de colocar conductores en paralelo, la resultante magnitud no habría sido aditiva. Las resistencias en serie se adicionan; en paralelo, no.

Las magnitudes que no son extensivas o aditivas se llaman *intensivas*. La misma operación de combinación de objetos puede dar lugar tanto a conceptos métricos extensivos como intensivos. Así, respecto a la operación de combinar dos economías nacionales para formar una

unión económica, los conceptos de *producto nacional bruto* o de *población* son extensivos o aditivos (pues el producto nacional bruto de la unión es igual a la suma de los productos nacionales y la población total es igual a la suma de las poblaciones), mientras que los conceptos de *renta per cápita* o de *tasa de natalidad* son intensivos. Respecto a la operación de vaciar el contenido de dos recipientes en un tercero el concepto de *volumen* es extensivo o aditivo, pero no los de *temperatura* o de *densidad*.

Consideremos con más detalle el caso de la temperatura como típico concepto métrico intensivo. Supongamos que ya disponemos de un concepto comparativo $<K, P>$ de temperatura en el dominio A (por ejemplo, el de los líquidos), basado en la prueba del tubo de mercurio sin graduar. Toda asignación t de números reales a los elementos de A que sea conforme con ese concepto comparativo (es decir, tal que si xKy, entonces $t(x)=t(y)$; y si xPy, entonces $t(x)< t(y)$) será una escala de temperatura. Pero, ¿cómo fijar una tal escala determinada? Aquí ya no basta con elegir un sólo objeto o proceso y asignarle un número (como en el caso de las magnitudes extensivas), sino que las cosas son algo más complicadas. En primer lugar, elegimos un cierto tipo de objetos del dominio A y asignamos un número c a estos objetos cuando se encuentra en un estado determinado y fácilmente reproducible. Luego asignamos otro número distinto k a los mismos objetos cuando se encuentran en otro estado determinado y fácilmente reproducible, pero distinto del anterior. Finalmente, introducimos una función de distancia d tal que para cada cuatro objetos x, y, z, w de A ocurre que si $d(x, y)=d(z, w)$, entonces $t(x)-t(y)=t(z)-t(w)$.

En el caso de la escala Celsius, lo que hacemos es asignar el número 0 al agua en el punto de fusión, el número 100 al agua en el punto de ebullición y elegir como función de distancia la distancia entre marcas hechas sobre el tubo de mercurio.

Escalas de intervalos

Sean f y h aplicaciones de un dominio A en **R**. Decimos que h es una transformación lineal positiva de f si y sólo si hay un número positivo r y un número cualquiera fijo s tales que para cada objeto x de A:

$$h(x)=r\cdot f(x)+s$$

Es decir, obtenemos el valor de h para x multiplicando el valor de f para x con un número positivo fijo y adicionando al resultado otro número determinado.

Decimos que un homomorfismo de un sistema empírico en otro numérico es una *escala de intervalos* si y sólo si toda transformación lineal positiva de ese homomorfismo es otro homomorfismo entre los mismos sistemas. Así como los conceptos métricos extensivos dan lugar a escalas proporcionales, los conceptos métricos intensivos dan lugar a escalas de intervalos. Las escalas de temperatura, por ejemplo, son escalas de intervalos. Por eso, para transformar una escala de temperatura en otra es necesario multiplicar por un número y adicionar otro. Así, para pasar de la escala centígrada a la escala Fahrenheit hemos de multiplicar el valor centígrado por $\frac{9}{5}$ y añadir 32 al resultado. Es decir:

$$t_F(x)=\frac{9}{5}t_C(x)+32.$$

Las escalas más informativas son las proporcionales. Todas las escalas proporcionales son automáticamente, además, escalas de intervalos y escalas ordinales. Las escalas de intervalos que no son proporcionales son algo menos ricas en información. Todas las escalas de intervalos son también escalas ordinales. Las escalas ordinales que no lo son de intervalos, las escalas meramente ordinales, son las más pobres en información. De hecho, como ya vimos, su rendimiento teórico no va más allá del de los conceptos comparativos.

Metrización fundamental y derivada

La metrización de un rasgo o característica de un ámbito determinado consiste en la introducción de un concepto métrico o magnitud para esa característica en ese ámbito determinado, o, dicho con más precisión, en el establecimiento de un homomorfismo (o clase de homomorfismos) entre el sistema empírico formado por dicho ámbito y dicha característica en un determinado sistema numérico.

En la práctica la metrización suele realizarse simplemente mediante una definición en función de otras magnitudes previamente introducidas. Así, podemos introducir el concepto métrico de *densidad* mediante la definición: densidad de x es igual a la masa de x dividida por el volumen de x, suponiendo que ya disponemos de los conceptos de masa y volumen. Igualmente podemos introducir el concepto métrico de *renta per cápita* mediante la definición: renta per cápita de x es igual al producto nacional de x dividido por la población de x, suponiendo que previamente hayamos introducido los conceptos de producto nacional y población.

Cuando introducimos un concepto métrico en función de otros previamente introducidos, decimos que se trata de una metrización derivada. La mayoría de las metrizaciones son derivadas. Por ejemplo, la introducción del concepto métrico de temperatura, anteriormente descrita, constituía una metrización derivada, pues utilizaba una función de distancia entre marcas, que presuponía el concepto métrico de longitud.

De todos modos, y aunque la mayoría de las magnitudes se introduzcan en función de otras, este procedimiento no puede seguirse con todas. Con alguno o con algunos conceptos métricos hay que empezar; alguna o algunas magnitudes han de ser introducidas sin presuponer la previa introducción de otras. En estos pocos, pero importantes casos, hablamos de metrización fundamental.

La introducción del concepto métrico de masa de que antes habíamos hablado constituye una metrización fundamental, pues no presuponía ninguna otra magnitud previa.

Los conceptos introducidos por metrización fundamental suelen referirse a ámbitos relativamente limitados. El concepto métrico de masa introducido fundamentalmente sólo era aplicable a los objetos físicos manejables. Pero no sólo queremos hablar de la masa de esos objetos. También queremos hablar de la masa de los átomos o de las estrellas, que no son manejables ni pueden colocarse en los platillos de una balanza. A este concepto generalizado de masa llegamos a través de una serie de hipótesis y teorías, de las que se desprende que el primer concepto de masa está correlacionado universalmente con otros conceptos de más amplio alcance, en función de los cuales podemos definir luego un nuevo concepto métrico de masa de más universal aplicabilidad.

En la génesis de muchos conceptos métricos importantes observamos esos dos momentos: la precisión de la idea intuitiva para un ámbito restringido, y la posterior ampliación de su alcance, redefiniéndolo en función de los nuevos conocimientos logrados.

Ventajas de los conceptos métricos

Las ventajas de los conceptos métricos respecto a los clasificatorios o comparativos son evidentes. El vocabulario cien-

tífico resulta mucho más simple, claro y manejable. Con un sólo concepto métrico tenemos infinitas posibles situaciones ya descritas y ordenadas, sin esfuerzo alguno de memoria. Si pretendiésemos sustituir un concepto métrico como el de temperatura por una serie de conceptos clasificatorios (gélido, frío, fresco, tibio, etc.), no sólo descendería considerablemente el nivel de precisión de nuestro lenguaje, sino que cargaríamos nuestra memoria con gran cantidad de términos distintos (y con su orden relativo).

Los conceptos métricos no sólo permiten formular leyes científicas mucho más sencillas y precisas que las formulables con términos cualitativos, sino que incluso tienen la ventaja heurística de facilitar la búsqueda de esas leyes. En efecto, si sospechamos una correlación entre dos magnitudes f y h, podemos medir los valores de f y h para diversos objetos o sucesos y, mediante un eje de coordenadas en el que los valores de f y de h estén marcados en los ejes de ordenadas y abscisas, respectivamente, señalar en el plano los puntos $< f(x_1), h(x_1) >$, $< f(x_2), h(x_2) >$, $< f(x_3), h(x_3) >$, etc. A continuación podemos trazar la curva más sencilla que pase por esos puntos y considerar la fórmula analítica que describa esa curva, como hipótesis. Posteriores mediciones confirmarán esa fórmula, o bien nos obligarán a trazar una curva más complicada, reformulando entonces la hipótesis, y así sucesivamente. De este modo se llega en algunos casos a la formulación de leyes científicas interesantes.

La razón profunda de todas las ventajas que se pueden aducir estriba en que los conceptos métricos constituyen un puente entre el mundo real y el mundo ideal de la matemática.

El mundo real de la naturaleza y la sociedad es un mundo en gran parte opaco a nuestra inteligencia, lleno de oscuros recovecos, siempre sorprendente, huidizo y poco manipulable intelectualmente.

El mundo de la matemática, por el contrario, es un mundo transparente, un mundo abierto a nuestra inteligencia, que lo ha creado y que lo abarca y manipula sin sorpresas, es un mundo perfectamente estructurado y ordenado y donde nos movemos con toda facilidad. Por eso, en cuanto los problemas que se plantean en el mundo real resultan demasiado complicados e inabarcables, la mejor estrategia para su solución suele consistir en representarlos como problemas relativos al mundo de la matemática, como problemas matemáticos, para los que ya sabemos cómo hallar una solución, solución que luego podemos retraducir al mundo real. Los conceptos métricos llevan a cabo esa representación del mundo real en el mundo de los números y nos permiten esa transposición de nuestras preguntas y de nuestros problemas sobre el mundo natural o social al mundo de la matemática, donde podemos usar todo el arsenal del cálculo diferencial e integral, del cálculo vectorial o tensorial, de la teoría de la probabilidad o la programación lineal, etcétera, para su solución y respuesta. Esta es la razón de que en general elijamos sistemas con el conjunto **R** de los números reales como sistemas numéricos en los que representar nuestros problemas. Aunque para realizar todas las medidas posibles e imaginables bastaría con los números racionales, la elección del conjunto de los números reales nos permite el uso de una artillería matemática más potente (como derivadas, integrales, ecuaciones diferenciales, etc.) para la resolución de nuestros problemas.

Es de esperar que una mejor comprensión de la estructura de la conceptualización científica no sólo sirva para subrayar una vez más, y por debajo de la confusa proliferación de las distintas especialidades, la profunda unidad de la empresa científica, sino que, además, contribuya a facilitar la introducción de nuevos y más precisos y fecundos conceptos en las áreas hasta ahora menos desarrolladas de la ciencia.

R: conjunto de los números reales.
R$^+$: conjunto de los números reales positivos.
f: A $\rightarrow$ B: f es una función que a cada elemento de A asigna un determinado elemento de B.

HOMOMORFISMOS

Sean $A \neq \emptyset$, $Q_1 \subset AXA$, $Q_2 \subset AXA$ y $g: AXA \rightarrow A$
Sean $S_1 \subset \mathbf{R}X\mathbf{R}$, $S_2 \subset \mathbf{R}X\mathbf{R}$ y $h: \mathbf{R}X\mathbf{R} \rightarrow \mathbf{R}$
Sea $\mathscr{A} = < A, Q_1, Q_2, g >$ y $\mathscr{R} = < \mathbf{R}, S_1, S_2, h >$

f es un homomorfismo de $\mathscr{A}$ en $\mathscr{R}$ $\Leftrightarrow$
$$\begin{cases} f: A \rightarrow \mathbf{R} \\ \wedge \quad \wedge\, xy \in A: (xQ_1y \Rightarrow f(x)\ S_1\ f(y)) \\ \wedge \quad \wedge\, xy \in A: (xQ_2y \Rightarrow f(x)\ S_2\ f(y)) \\ \wedge \quad \wedge\, xyz \in A: (g(x, y) = z \Rightarrow h\,(f(x), f(y)) = f(z)) \end{cases}$$

TRANSFORMACIONES

Sean f: A $\rightarrow$ **R** y h: A $\rightarrow$ **R**

h es una transformación monótona de $f \Leftrightarrow \wedge\, xy \in A: (f(x) < f(y) \Rightarrow h(x) < h(y))$
h es una transformación lineal positiva de $f \Leftrightarrow \vee\, r \in \mathbf{R}^+\ \vee s \in \mathbf{R} \wedge x \in A: h(x) = r \cdot f(x) + s$
h es una transformación similar de $f \Leftrightarrow \vee\, r \in \mathbf{R}^+ \wedge x \in A: h(x) = r \cdot f(x)$

ESCALAS

f es una escala ordinal de $\mathscr{A}$ en $\mathscr{R}$ $\Leftrightarrow$
$$\begin{cases} f \text{ es un homomorfismo de } \mathscr{A} \text{ en } \mathscr{R} \\ \wedge \quad \wedge\, h\,(h \text{ es una transformación monótona de } f \Rightarrow \\ \Rightarrow h \text{ es un homomorfismo de } \mathscr{A} \text{ en } \mathscr{R}) \end{cases}$$

f es una escala de intervalos de $\mathscr{A}$ en $\mathscr{R}$ $\Leftrightarrow$
$$\begin{cases} f \text{ es un homomorfismo de } \mathscr{A} \text{ en } \mathscr{R} \\ \wedge \quad \wedge\, h\,(h \text{ es una transformación lineal positiva de } f \Rightarrow \\ \Rightarrow h \text{ es un homomorfismo de } \mathscr{A} \text{ en } \mathscr{R}) \end{cases}$$

f es una escala proporcional de $\mathscr{A}$ en $\mathscr{R}$ $\Leftrightarrow$
$$\begin{cases} f \text{ es un homomorfismo de } \mathscr{A} \text{ en } \mathscr{R} \\ \wedge \quad \wedge\, h\,(h \text{ es una transformación similar de } f \Rightarrow \\ \Rightarrow h \text{ es un homomorfismo de } \mathscr{A} \text{ en } \mathscr{R}) \end{cases}$$

UNA ESCALA ES UN HOMOMORFISMO de un sistema empírico $\mathscr{A}$ en un sistema numérico $\mathscr{R}$. El conjunto de las transformaciones del mismo tipo de esa escala constituye un concepto métrico. Y el rendimiento teórico del concepto métrico es tanto mayor cuanto más exigente es el tipo de transformación permitido. Aquí definimos las nociones de homomorfismo, transformación y escala usando los signos matemáticos introducidos al principio de este cuadro y del de la página 83.

Juegos matemáticos

Donde al unir mediante líneas los puntos de un conjunto se recorren caminos diversos, divergentes y divertidos

Martin Gardner

"Demostrar que en toda reunión de seis personas habrá tres que se conozcan mutuamente o que se sean totalmente extrañas".

– Problema E 1321.
The American Mathematical Monthly,
junio-julio, 1958

Mi sección de este mes, dedicada a la teoría de grafos de Ramsey, es homenaje a *The Journal of Graph Theory,* revista aparecida en 1977 y dedicada exclusivamente a la teoría de grafos, una de las ramas de la moderna matemática que se está desarrollando a mayor velocidad. Su editor es Frank Harary, eminente teórico, autor del texto más difundido sobre la materia. La revista es publicada por John Wiley & Sons.

La teoría de grafos estudia conjuntos de puntos unidos mediante líneas. Dos artículos del primer número de la nueva revista trataban de teoría de grafos de Ramsey, tema que se solapa ampliamente con la matemática recreativa. Aunque hacia la década de 1930 se publicaron unos cuantos artículos sobre la teoría de Ramsey, debidos al matemático húngaro Paul Erdös y otros autores, la labor de investigación de los llamados números de Ramsey no comenzó a desarrollarse en serio hasta finales de la década de 1950. Uno de los grandes estímulos para realizar esta investigación fue el problema, de inocente aspecto, arriba anunciado.

Es fácil transformar este problema en otro de teoría de grafos. Las seis personas se representan mediante seis puntos. Se une cada par de puntos mediante una línea, usando, por ejemplo, lápiz rojo para indicar que dos personas se conocen, y lápiz azul para mostrar que se son extrañas. El problema consiste ahora en demostrar que, no importa cuál sea el color que se asigne a estas líneas, se producirá inevitablemente un triángulo rojo (cuyos lados conectan tres personas que se conocen mutuamente) o un triángulo azul (que une a tres mutuamente desconocidas).

La teoría de Ramsey, que estudia este tipo de problemas, recibe su nombre de un extraordinario matemático de la Universidad de Cambridge, Frank Plumpton Ramsey. Contaba solamente 26 años al morir, en 1930, pocos días después de sufrir una operación abdominal. Su padre era presidente del Colegio de la Magdalena, en Cambridge, y su hermano fue arzobispo de Canterbury desde 1961 hasta 1974. Entre los economistas, Ramsey es conocido por sus valiosas contribuciones a la teoría económica. Entre los especialistas en lógica su fama se debe a la simplificación realizada por él de la teoría ramificada de tipos, de Bertrand Russell (se dice que Ramsey "ramsificó" la teoría ramificada), así como por su clasificación de las paradojas lógicas en paradojas estrictamente lógicas y paradojas de naturaleza semántica. En la filosofía de la ciencia se le conoce por haber interpretado la probabilidad en términos de creencias, y por la invención de la "sentencia de Ramsey": un mecanismo simbólico que aclara notablemente la naturaleza del "lenguaje teórico" de la ciencia.

En 1928, Ramsey leyó ante la London Mathematical Society su ahora clásico artículo "On a Problem in Formal Logic". (Puede encontrarse una reimpresión en *The Foundations of Mathematics,* recopilación póstuma de ensayos de Ramsey, realizada por R. B. Baithwaite, y ahora disponible en rústica.) En dicho artículo Ramsey demostró un profundo resultado de teoría de conjuntos, hoy conocido por teorema de Ramsey. Empezó demostrándolo para conjuntos infinitos, al observar que la demostración era en este caso más sencilla que en el de conjuntos finitos, caso que probó a continuación. Lo mismo que otros muchos teoremas conjuntistas, el de Ramsey resultó tener gran variedad de insospechadas aplicaciones en problemas combinatorios. El teorema es, en toda su fuerza y generalidad, demasiado complicado para poder presentarlo aquí en esa forma; para nuestros propósitos es suficiente ver cómo puede aplicarse a la teoría de coloración de grafos.

Si se une, mediante una línea, todo par de puntos de un conjunto de n puntos se obtiene un grafo completo de n puntos, que se representa por K_n. Como tan sólo nos interesan las propiedades topológicas del grafo, no importa cómo se sitúen los puntos o cómo se tracen las líneas. La ilustración de la página 96 muestra la forma ordinaria de dibujar los grafos completos correspondientes a conjuntos de dos a seis puntos. Las líneas identifican cada uno de los subconjuntos de n puntos que tienen exactamente dos elementos.

Supongamos que las líneas de un grafo K_n se colorean arbitrariamente de rojo o de azul. Podríamos pintarlas todas de rojo, o todas de azul, o repartirlas entre ambos colores en la forma que más nos plazca. Al hacerlo así se obtiene un bicoloreado del grafo. Evidentemente, el coloreado no es sino un método sencillo de separar todos los subconjuntos de dos elementos en dos clases mutuamente excluyentes. Análogamente, usando tres colores las líneas se dividen en tres clases. De forma general, un coloreado de las líneas que utilice r colores (un r-coloreado) descompone al conjunto de pares de puntos en r clases mutuamente excluyentes.

Un "subgrafo" de un grafo completo es cualquier tipo de grafo contenido en el grafo completo, contenido en el sentido de que todos los puntos y líneas del subgrafo pertenecen al grafo mayor. Es fácil comprobar que todo grafo completo es subgrafo de todo grafo completo de más puntos. Muchos grafos sencillos han recibido nombres especiales. La ilustración de la página 97 muestra cuatro familias: caminos, ciclos, estrellas y ruedas. Nótese que la rueda correspondiente a cuatro puntos no es sino otra forma de dibujar K_4, y suele llamarse "tetraedro" porque es una proyección planar del esqueleto de un tetraedro.

Consideremos ahora el siguiente problema, para el que se necesitan seis lápices de distintos colores. Asociemos con cada color un tipo de grafo cualquiera, como más nos agrade. Por ejemplo:

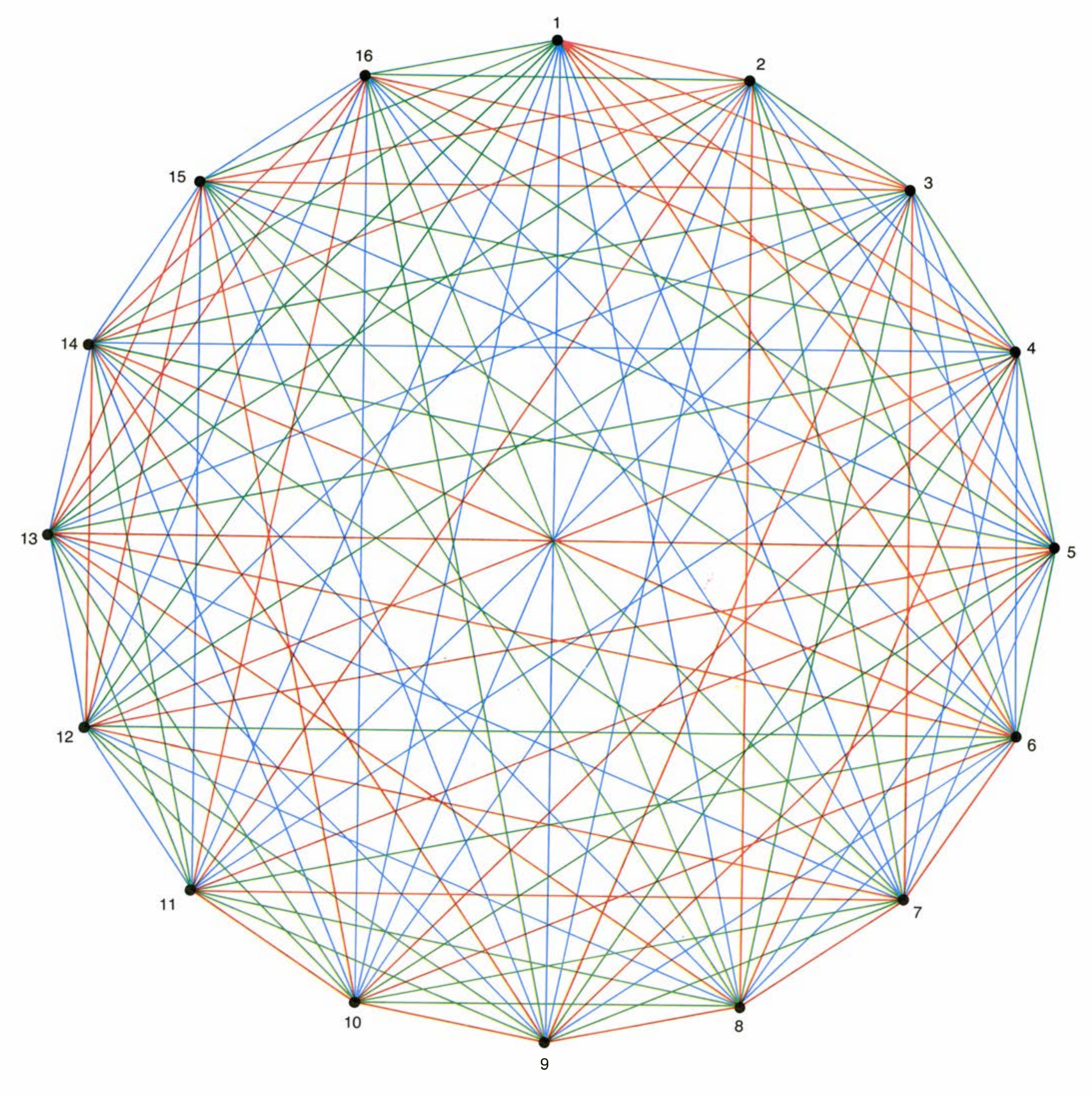
1. Rojo: un pentágono (ciclo de cinco puntos).
2. Naranja: un tetraedro.
3. Amarillo: una estrella de siete puntos.

4. Verde: un camino de 13 puntos.
5. Azul: una rueda de ocho puntos.
6. Púrpura: un lazo de pajarita (dos triángulos que comparten solamente un punto).

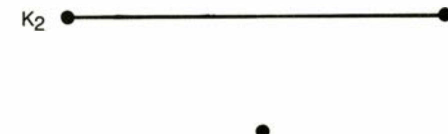
Planteemos ahora una curiosa pregunta: ¿Existen grafos completos que, coloreados arbitrariamente con estos seis colores, contengan con certeza como subgrafo al menos uno de los seis grafos arriba enumerados? Dicho de otra forma, se trata de hallar un grafo completo que coloreado con estos seis lápices –no importa cómo– permita asegurar que con certeza se obtendrá un pentágono rojo, o un tetraedro naranja, o una estrella de siete puntos amarilla, etcétera. El teorema de Ramsey prueba que, a partir de cierto tamaño finito, todo grafo completo posee esta propiedad. El grafo mínimo de este conjunto infinito se llama grafo de Ramsey del conjunto de grafos especificado, y el número de vértices que contenga es el número de Ramsey de dicho conjunto de subgrafos.

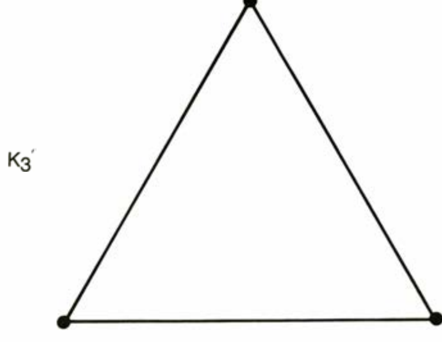
Todo grafo de Ramsey origina un juego y plantea un problema. En nuestro ejemplo el juego consistirá en lo siguiente. Dos jugadores eligen alternativamente uno cualquiera de los seis lápices y trazan con él un lado en el grafo de Ramsey. El primero de ellos que complete el trazado de uno de los seis subgrafos dados pierde el juego. Por tratarse de un grafo de Ramsey, la partida no puede terminar en tablas. El grafo de Ramsey es, además, el mínimo grafo completo que posee esta propiedad.

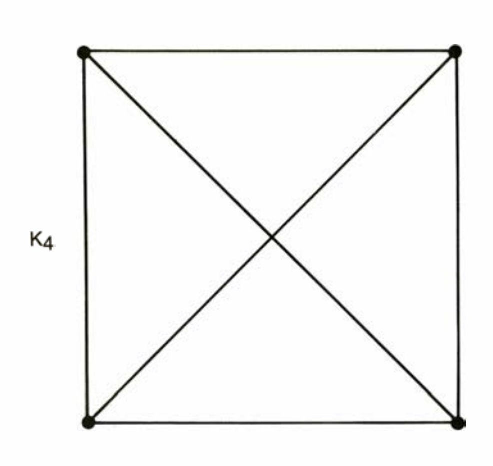
El correspondiente problema hace uso de un grafo completo con un punto menos que el de Ramsey. Se trata, evidente-

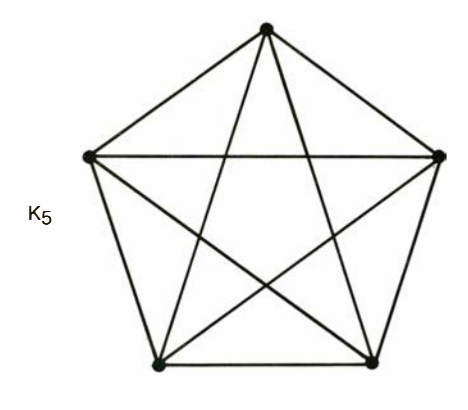


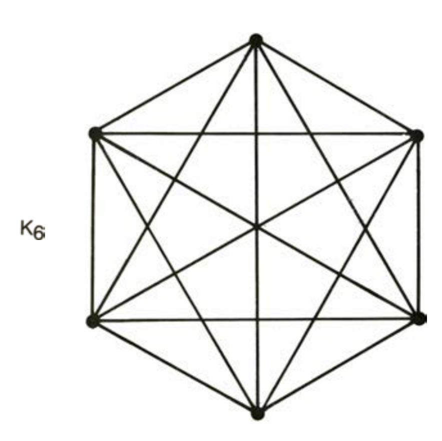
Un tricoloreado crítico del grafo completo K_{16}, de 16 puntos











Grafos completos de 2 a 6 puntos

mente, del máximo grafo completo donde el juego puede terminar en tablas. Tal grafo se llama grafo de Ramsey crítico para el conjunto de subgrafos especificado. El problema consiste en hallar una coloración del grafo crítico con la que no llegue a formarse ninguno de los grafos dados. Una tal coloración se llama coloración crítica.

No tengo la menor idea de cuánto pueda valer el número de Ramsey correspondiente a los seis subgrafos dados. Su grafo completo sería tan grande –seguramente contendrá centenares de puntos– que no hay que tomar en cuenta la posibilidad de jugar una partida con él; y el correspondiente problema es tan difícil que ni siquiera puede considerarse factible la búsqueda de soluciones mediante computador. No obstante, los juegos de Ramsey y sus correspondientes rompecabezas pueden resultar muy entretenidos si se utilizan grafos completos más pequeños y lápices de tan sólo dos colores.

El más conocido de los juegos de Ramsey, llamado Sim, fue examinado en esta sección de *Scientific American* en enero de 1973. Se juega con el grafo completo de seis puntos (K_6), que sirve de modelo al problema de la reunión de seis personas. No es difícil demostrar que el número de Ramsey de los dos subgrafos siguientes es igual a seis:

1. Rojo: un triángulo (K_3).
2. Azul: un triángulo (K_3).

En la teoría de Ramsey "clásica" es costumbre usar números solos para denotar grafos completos, pudiéndose así expresar el resultado anterior con la siguiente notación abreviada: $R(3,3)=6$, lo que significa que R, número de Ramsey del mínimo grafo completo que forzosamente contiene un triángulo "monocromático" (completamente rojo o completamente azul), al colorearlo con dos tonos, es igual a seis. Así pues, si dos jugadores van alternativamente coloreando el grafo K_6 de rojo y azul, es seguro que un jugador habrá de perder, viéndose obligado a colorear todo un triángulo de su color. El correspondiente (fácil) problema consiste en pintar de dos colores el grafo crítico K_5 de forma que no resulte ningún triángulo monocromático.

Al pintar K_6 de dos colores se forman necesariamente, cuando menos, dos triángulos monocromáticos. (Si hay exactamente dos que sean de distintos colores formarán un lazo de pajarita.) Se plantea así una interesante cuestión. Si se pinta de dos colores un grafo completo de n puntos, ¿cuántos triángulos monocromáticos resultarán forzosamente? El primero en responder a ella fue A. W. Goodman, en un artículo titulado "On Sets of Acquaintances and Strangers at Any Party" ("Sobre los conjuntos de personas mutuamente conocidas y desconocidas de una reunión"), publicado en *The American Mathematical Monthly* (vol. 66, n.º 9, noviembre de 1959, páginas 778-783). Para presentar la fórmula de Goodman lo mejor es descomponerla en tres casos:

Si n es de la forma $2u$, el número de triángulos monocromáticos que forzosamente habrán de aparecer será $(^1/_3)u(u-1)(u-2)$. Si n es de la forma $4u+1$, este número resulta ser $(^1/_3)2u(u-1)(4u+1)$. Si n es $4u+3$, resulta ser $(^1/_3)2u(u-1)(4u-1)$. Así pues, para grafos completos de seis a 12 puntos el número de triángulos forzosos de un solo color es, respectivamente, dos, cuatro, ocho, 12, 20, 28 y 40.

Si se realiza al azar el coloreado de los lados se producirán, por lo común, más triángulos monocromáticos de los obligados. Cuando el coloreado de un grafo de Ramsey contiene exactamente el número de triángulos forzosos, y ninguno más, se dice que el coloreado es extremal. ¿Existe siempre un coloreado extremal en el que todos los triángulos forzosos sean del mismo color? (Tales coloreados se han llamado vacíos de azul, con lo que se pretende decir que el número de triángulos azules se ha reducido a cero). En 1961, Léopold Sauvé demostró que la respuesta es afirmativa, lo que induce a pensar en una nueva clase de rompecabezas. Por ejemplo, tracemos el grafo completo de siete puntos. ¿Podrá el lector colorearlo de manera que no haya triángulos azules y tampoco resulten más de cuatro triángulos rojos? No es muy fácil conseguirlo. (El mes próximo se dará una solución.)

Se sabe muy poco acerca de los números de Ramsey "clásicos", que son el número de puntos del mínimo grafo completo en el que forzosamente estará contenido un conjunto dado de grafos completos menores. No se conoce ningún procedimiento práctico para determinar números de Ramsey clásicos. Se conoce un algoritmo: sencillamente, se exploran todos los posibles métodos de colorear grafos completos, de orden cada vez mayor, hasta encontrar el grafo de Ramsey. La dificultad de la tarea crece exponencialmente a tal velocidad que pronto resulta impracticable incluso la exploración computarizada. Todavía se sabe menos acerca del vencedor –si será el primer jugador o el segundo– en el caso de

jugarse racionalmente una partida de un juego de Ramsey. El juego llamado Sim ha sido resuelto (siempre puede vencer el segundo jugador), pero no se sabe prácticamente nada sobre los juegos de Ramsey con grafos completos mayores.

Hasta el momento nos hemos limitado a considerar el tipo de juegos de Ramsey que Harary llama "juegos de evitación". Como él mismo ha señalado, son posibles al menos otras tres clases de juego. Por ejemplo, en un juego de "consecución" (que se desarrolla de la misma forma que el Sim), el primer jugador que consiga completar un triángulo gana la partida. En los otros dos juegos, la partida prosigue hasta que todas las líneas se hayan coloreado, ganando entonces bien el jugador que tenga más triángulos de su color o bien el que tenga menos. Estos dos últimos juegos son los de más difícil análisis, resultando el de "consecución" el más sencillo. En lo sucesivo, "juego de Ramsey" significará juego de evitación.

Además de $R(3,3)$, que es base del Sim, se conocen otros cinco números de Ramsey clásicos, no triviales, para coloreados de dos tonos:

1. $R(3,4)$ es igual a 9. Si K_9 se pinta de dos colores resulta forzosamente un triángulo rojo (K_3) o un tetraedro azul (K_4). No se conoce quién es vencedor en el juego de Ramsey.
2. $R(3,5)$ es igual a 14.
3. $R(4,4)$ es igual a 18. Si K_{18} se pinta de dos colores, resulta un tetraedro monocromático (K_4) forzosamente. El correspondiente juego de Ramsey no está mal, aunque no sea cómodo jugarlo por la dificultad de identificar tetraedros. El grafo y su coloración ponen de manifiesto que en una reunión de 18 personas hay un conjunto de cuatro que se conocen mutuamente o que son mutuamente desconocidas.
4. $R(3,6)$ es igual a 18. En la reunión del ejemplo anterior o bien hay un conjunto de tres personas que se conocen entre sí, o bien hay un conjunto de seis que se son totalmente extrañas. La forma en que estos dos conjuntos están relacionados con los dos conjuntos tetraédricos anteriores es una interesante cuestión que, según parece, nadie ha investigado todavía.
5. $R(3,7)$ es igual a 23.

Obsérvese que en la lista anterior no figura $R(5,5)$. Ello es debido a que no se conoce todavía el número de Ramsey de un grafo completo que, pintado de dos colores, contenga necesariamente un K_5 monocromático. Stephen A. Burr, del departamento de líneas de larga distancia de American Telephone and Telegraph Company, y distinguido especialista en teoría de grafos de Ramsey, opina que posiblemente $R(5,5)$ no llegue a conocerse nunca; tan grande es el salto en aumento de complejidad. Incluso $R(4,5)$ es tan difícil que, según cree, concebiblemente no llegue tampoco a ser hallado jamás. Sin embargo, en ambos casos se conocen acotaciones. $R(5,5)$ se encuentra entre 38 y 67, ambos inclusive; $R(4,5)$ está entre 25 y 29, ambos inclusive.

Solamente se conoce otro número de Ramsey clásico, y es para tres colores. $R(3)=3$, como es evidente; pues si se colorea un triángulo con un solo color es seguro que se obtendrá un triángulo monocromático. Hemos dicho ya que $R(3,3)$ es igual a 6. $R(3,3,3)$ es igual a 17, lo que significa que si K_{17} se pinta de tres colores, necesariamente se habrá de obtener un triángulo monocromático. En realidad, resultará más de uno de tales triángulos, pero no se conoce su número exacto.

La demostración de que $R(3,3,3)=17$ se obtuvo por primera vez en 1955. En el juego de Ramsey correspondiente a este grafo se usan lápices de tres colores. Los jugadores van trazando líneas alternativamente, del color que prefieran, hasta que uno de ellos, viéndose forzado a completar un triángulo monocromático, pierde. ¿Quién será el vencedor, si ambos jugadores realizan siempre movimientos óptimos? Nadie lo sabe. El correspondiente problema de Ramsey consiste en tricolorear el grafo crítico K_{16} sin que aparezca ningún triángulo monocromático. La ilustración de la página 95, que se ha reproducido de *Graphs and Hypergraphs,* obra debida a Claude Berge, muestra una de las dos soluciones esencialmente distintas conocidas. ("Distintas" recibe aquí un sentido combinatorio más profundo que la mera supresión de giros y simetrías.)

PUNTOS	CAMINOS	CICLOS	ESTRELLAS	RUEDAS
2				
3				
4				
5				
6				

Cuatro familias importantes de grafos simples

GRAFO	NUMERO RAMSEY	GRAFO	NUMERO RAMSEY	GRAFO	NUMERO RAMSEY	GRAFO	NUMERO RAMSEY	GRAFO	NUMERO RAMSEY	GRAFO	NUMERO RAMSEY
1	2	20	9	39	9	58	11	77	9	96	10
2	3	21	8	40	9	59	11	78	9	97	11
3	6	22	7	41	11	60	11	79	9	98	10
4	5	23	8	42	10	61	11	80	9	99	11
5	5	24	7	43	11	62	8	81	10	100	11
6	6	25	8	44	10	63	10	82	9	101	11
7	6	26	7	45	12	64	11	83	10	102	13
8	7	27	7	46	14	65	10	84	9	103	12
9	10	28	8	47	9	66	11	85	10	104	12
10	18	29	8	48	10	67	11	86	9	105	11
11	6	30	8	49	9	68	9	87	9	106	12
12	7	31	10	50	10	69	9	88	9	107	11
13	6	32	8	51	7	70	9	89	11	108	11
14	7	33	8	52	10	71	9	90	10	109	14
15	6	34	10	53	8	72	9	91	11	110	13
16	9	35	9	54	11	73	9	92	11	111	13
17	6	36	8	55	8	74	10	93	10	112	15
18	9	37	9	56	11	75	9	94	11	113	17
19	9	38	9	57	8	76	9	95	10		

Grafos simples cuyo número de Ramsey generalizado es conocido

¿Qué puede decirse de $R(3,3,3,3)$, que es el mínimo grafo completo que obliga a la presencia forzosa de un triángulo monocromático al pintarlo de cuatro colores? Todavía no se sabe, aunque Jon Folkman, brillante especialista de matemática combinatoria, demostró que sería menor o igual que 64. Folkman se suicidó en 1964, cuando contaba 31 años, tras sufrir una operación quirúrgica motivada por un enorme tumor cerebral. La mejor de las cotas inferiores conocidas es 51, y fue determinada por Fan Chung, joven matemático chino de los Bell Laboratories, quien presentó la demostración en su tesis doctoral.

La teoría clásica de Ramsey puede generalizarse en multitud de formas fascinantes. Hemos considerado ya la más evidente: la búsqueda de los llamados números de Ramsey generalizados para *r*-coloreados de grafos completos que obligan a la aparición forzosa de subgrafos distintos de los grafos completos. Váćlav Chvátal y Harary fueron pioneros en este territorio, y Burr ha estado explotándolo durante los cinco últimos años. Consideremos el problema de hallar los números de Ramsey de los grafos completos mínimos que implican la aparición de una estrella monocromática de *n* puntos. Harary fue el primero en resolverlo para coloreados bicromáticos. En 1973, Burr y J. A. Roberts lo resolvieron para cualquier número de colores.

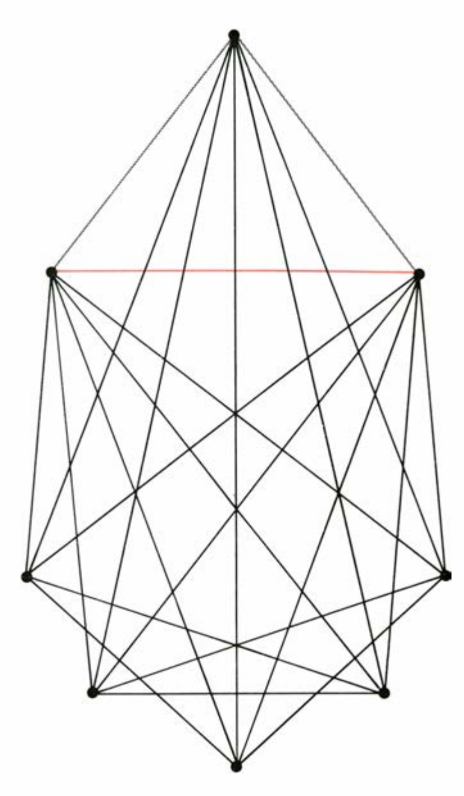
Otro problema de Ramsey generalizado consiste en hallar los números de Ramsey para coloraciones bicromáticas de K_n que impliquen un número prefijado de triángulos monocromáticos "disjuntos". (Dos triángulos son disjuntos si no tienen ningún punto común.) En 1975, Burr, Erdös y J. H. Spencer demostraron que tal número es $5d$, siendo d(mayor que dos) el número de triángulos disjuntos. El problema está sin resolver para más de dos colores.

Para ruedas, el caso general no está resuelto ni siquiera con dos colores. El número de Ramsey para la rueda de cuatro puntos, o tetraedro, es, como se ha visto, igual a 18. Se ha demostrado recientemente que el número de Ramsey de la rueda de cinco puntos (una rueda con un eje y cuatro radios) es igual a 15. La demostración se debe al matemático nigeriano Tim Moon. El problema está sin resolver para la rueda de seis puntos, aunque se sabe que su número de Ramsey está comprendido entre 17 y 20, ambos inclusive. Se ha conjeturado que su valor sea 20. Si así fuese, se podrían colorear los grafos completos de 17, 18 y 19 puntos sin que tuviera que aparecer forzosamente una rueda monocromática de seis puntos (cinco radios). Tendría considerable interés poder obtener una tal coloración de K_{17}, y tendré sumo placer en recibir noticias de aquellos lectores que encuentren una. Caso de existir, probablemente presente alto grado de simetría.

En la ilustración de la página anterior se publica aquí por primera vez una valiosa carta preparada por Burr, que relaciona los 113 grafos simples de no más de seis líneas y sin puntos aislados para los cuales se conoce el número de Ramsey generalizado. Obsérvese que algunos de estos grafos no son conexos. En tales casos toda la configuración, bien totalmente de rojo, bien en azul, viene implicada por el grafo completo cuyo número de Ramsey se indica.

Cada grafo de la carta de Burr sirve de base a un juego de Ramsey y a un rompecabezas, si bien estos últimos –consistentes en hallar coloreados críticos de los grafos críticos– resultan mucho más fáciles de hallar que los coloreados críticos para los números de Ramsey clásicos. Obsérvese que la carta permite seis variantes del juego llamado Sim. Una coloración bicromática de K_6 no solamente implica la aparición de un triángulo monocromático, sino también un cuadrado, una estrella de cuatro puntos (llamada a veces una "garra"), un camino de cinco puntos, un par de caminos disjuntos de dos y tres puntos (ambos del mismo color), un cuadrado con una cola, y el "árbol" simple que ocupa el 15-avo lugar en la carta de Burr. El triángulo con cola (8), la estrella de cinco puntos (12), la cruz latina (27) y el pez (51), seguramente ofrecerán interesantes juegos de Ramsey en K_7.

Ronald L. Graham, uno de los más destacados especialistas en matemática combinatoria de los Estados Unidos, y jefe del Departamento de Matemática Finita de los Bell Laboratories, ha efectuado muchas y valiosas aportaciones a la teoría de Ramsey generalizada. Verdaderamente, sería muy difícil hallar un matemático creador que se ajuste menos al estereotipo cinematográfico correspondiente. En su temprana juventud, Graham y dos de sus amigos actuaron como acróbatas profesionales de trampolín, trabajando en un circo con el nombre artístico de "Bouncing Baers". Es también uno de los mejores malabaristas estadounidenses, y ha sido presidente de la International Juggler's Association (Asociación Internacional de Malabaristas). El techo de su oficina está cubierto con una gran red, que puede bajarse y sujetarse a la cintura. De esta forma, cuando practica con seis o siete bolas, las que pueda fallar retornan servicialmente hasta él.



Solución a un problema de Erdös

Graham halló en 1968 una ingeniosa solución de un problema del tipo de Ramsey, propuesto por Erdös y András Hajnal. ¿Cuál es el grafo mínimo, no importa de qué tipo, que no contiene a K_6 y que pintado de dos colores implica la aparición de un triángulo monocromático? La única solución, dada por Graham, es el grafo de ocho puntos que se muestra en la ilustración superior de esta página. La demostración se obtiene por reducción al absurdo. Se comienza suponiendo posible una coloración bicromática que no forme ningún triángulo monocromático y se demuestra después que esta hipótesis implica la existencia de uno de tales triángulos. Al menos dos líneas de las que parten del vértice más alto han de ser, por ejemplo, grises, y por la simetría del grafo podemos hacer que lo sean las dos más exteriores sin por ello perder generalidad. Los extremos de estas dos líneas habrán de unirse por una línea de color para impedir que se forme un triángulo gris. Quizá les resulte interesante a los lectores completar el razonamiento.

¿Qué puede decirse de otros problemas parecidos en los que el subgrafo excluido sea un grafo completo pero distinto de K_6? La cuestión carece de sentido para K_3, porque el propio K_3 es un triángulo. Para K_5 el problema dista de estar completamente resuelto. Se tiene un gra-

fo de 18 puntos, publicado por Robert W. Irving en un artículo de 1973 que soluciona parcialmente el problema. Mucho más lejos de la solución está el caso K_4. Folkman, en un artículo publicado después de su muerte, demostró que tal grafo de Ramsey existe, pero en su construcción empleaba más de $2\uparrow\uparrow\uparrow2^{901}$ puntos. Número tan monstruosamente grande, que no existe forma de escribirlo sin emplear la notación especial, de flechas, anterior. Esta notación sagital fue introducida por Donald E. Knuth en su artículo "Mathematics and Computer Science: Coping with Finiteness", en *Science* (17 de diciembre de 1976).

Imaginemos el universo atiborrado de esferillas del tamaño de un electrón. El número total de estas esferas sería inconcebiblemente menor que el número de puntos del grafo de Folkman. Erdös tiene en vigor una oferta de 100 dólares a quien encuentre un grafo que solucione el problema con menos de un millón de puntos.

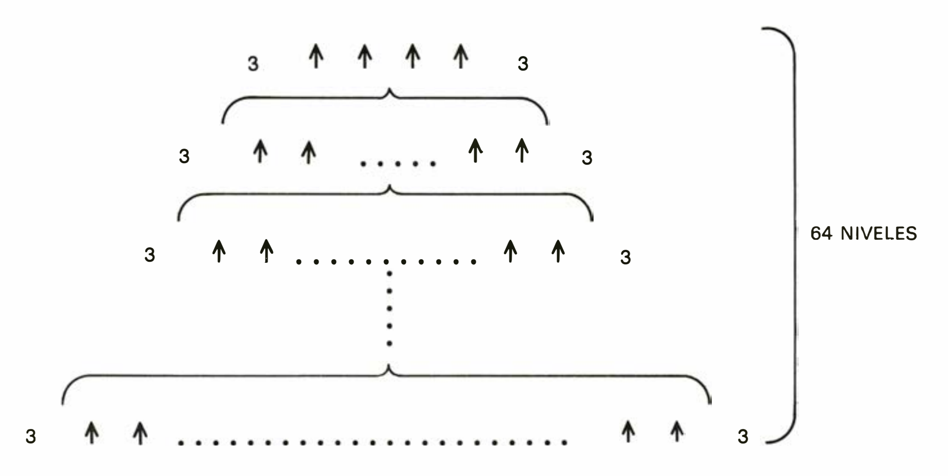
El grafo de Folkman muestra con tremenda claridad cuán difícil puede ser un problema de Ramsey incluso cuando en su enunciado no se mencionen grafos de más de cuatro puntos. Pero como solía decir Al Jolson, "eso no es ná de ná". Graham ha encontrado un ejemplo que puede causar mayor sobresalto todavía.

Consideremos un cubo en el que todo par de vértices esté unido por una recta. Se tiene así un grafo completo de ocho puntos, aunque ahora también se tendrá en cuenta la estructura geomátrica euclidiana. Imaginemos que las líneas de este K_8 tridimensional se colorean arbitrariamente de rojo y de azul. ¿Es posible evitar que aparezca un K_4 monocromático *contenido en un plano*? No es difícil comprobar que así ocurre.

Generalicemos el problema a cubos de dimensión n. Un hipercubo tiene 2^n vértices. En el hipercubo de cuatro dimensiones también pueden pintarse de dos colores las líneas del grafo completo de $2^4=16$ puntos de forma que no aparezca ningún grafo planar monocromático de cuatro puntos. Lo mismo puede hacerse para el hipercubo de cinco dimensiones y $2^5=32$ vértices. Estos ejemplos sugieren este problema euclídeo de Ramsey: ¿Cuál es el número mínimo de dimensiones que ha de tener un hipercubo para que al pintar usando dos colores las líneas que unen todo par de vértices forzosamente aparezca un K_4 planar pintado de un solo color? El teorema de Ramsey solamente garantiza que el problema tenga solución si no se impone la restricción de que el K_4 de forzosa aparición haya de estar contenido en un plano.

La existencia de solución cuando el K_4 obligatorio haya de ser planar fue demostrada por primera vez por Graham y Bruce L. Rothschild en una generalización de gran alcance del teorema de Ramsey que ellos descubrieron en 1970. Ahora bien, la determinación efectiva de la dimensión perdida ya es otra cuestión. Graham, en una demostración todavía no publicada, ha obtenido recientemente una cota superior, pero ésta es tan enorme que posee el récord de ser el máximo número que jamás se haya utilizado en una demostración matemática seria.

Para captar, aunque sea muy someramente, el tamaño del número de Graham, hemos de explicar primero la notación sagital de Knuth. El número escrito $3\uparrow3$ es $3\times3\times3=3^3=27$. El número $3\uparrow\uparrow3$ denota la expresión $3\uparrow(3\uparrow3)$. Como $3\uparrow3$ es igual a 27, podemos expresar $3\uparrow\uparrow3$ en la forma $3\uparrow27$ o también, 3^{27}. En forma de torre inclinada de exponentes sería:



La cota superior de Graham para la solución del problema de Ramsey euclídeo

La torre tiene solamente tres pisos de altura, pero expresada en la forma ordinaria es 7.625.597.484.987. Se ha dado ya un gran salto desde 27, pero todavía es un número pequeño que es posible imprimir sin dificultad.

Cuando el enorme número $3\uparrow\uparrow\uparrow3=3\uparrow\uparrow(3\uparrow\uparrow3)=3\uparrow\uparrow3^{27}$ se escribe en forma de torre de 3, alcanza una altura de 7.625.597.484.987 pisos. Tanto la torre, como el número por ella representado son ahora demasiado enormes para poder imprimirlos sin recurrir a notaciones especiales.

Consideremos ahora $3\uparrow\uparrow\uparrow\uparrow3=3\uparrow\uparrow\uparrow(3\uparrow\uparrow\uparrow3)$. Dentro del paréntesis se encuentra el altísimo número obtenido en el cálculo precedente. Ya no es posible expresar de manera sencilla la altura de la torre de treses que expresa $3\uparrow\uparrow\uparrow\uparrow3$. Su altura está a un universo de distancia de $3\uparrow\uparrow\uparrow3$. Si se descompone $3\uparrow\uparrow\uparrow\uparrow3$ en una sucesión de operaciones sagitales dobles resulta $3\uparrow\uparrow(3\uparrow\uparrow(3\uparrow\uparrow...\uparrow\uparrow(3\uparrow\uparrow3)...))$, donde el número de pasos a iterar es $3\uparrow\uparrow\uparrow3$. Como dice Knuth con cierta ironía, "los puntos suprimen muchos detalles", $3\uparrow\uparrow\uparrow\uparrow3$ es inimaginablemente mayor que $3\uparrow\uparrow\uparrow3$, pero todavía es pequeño, pues, como es evidente, la mayoría de los números finitos son mucho mayores que él.

Estamos ahora preparados para indicar el número de Graham. Está representado en la ilustración inferior. En la parte más alta se encuentra $3\uparrow\uparrow\uparrow\uparrow3$, que indica el número de flechas del número inmediatamente inferior. A su vez, este número expresa el número de flechas del inferior, y así sucesivamente durante 2^6, es decir, 64, capas. El número final es el número que Graham ha demostrado ser cota superior de la solución del problema del hipercubo.

Pues ahora agárrese al sombrero, los expertos en teoría de Ramsey opinan que el verdadero número de Ramsey de este problema es probablemente igual a 6. Como Stanislav M. Ulam ha dicho muchas veces en sus lecciones: "Con el infinito nos defenderemos bien. Lo finito puede costarnos algún tiempo más".

El problema del mes pasado consistía en adivinar la regla secreta que determinaba la disposición final de las cartas en una mano del juego llamado "Eleusis". La regla era: "Si la última carta es menor que la precedente lícitamente jugada, jugar una carta más alta que la última carta; en caso de que ocurra lo contrario, jugar una más baja. La primera carta es correcta salvo si es igual a la carta de apertura".

Taller y laboratorio

Maravillas de la física que pueden hallarse en una taza de café o de té

Jearl Walker

En una época en la cual la mayoría de los hallazgos de la ciencia se dan en cosas que son totalmente extrañas a nuestro quehacer diario –las partículas fundamentales de la materia, la actividad que se desarrolla en el interior de las células vivas o el singular comportamiento de lejanos objetos celestes–, se tiende a no prestar atención a las oportunidades de estudiar hechos científicos que se producen delante de nuestra mirada. Este mes he reunido varios ejemplos, basados en una cosa tan sencilla como una taza de café o de té. Sin duda, el lector se habrá encontrado con algunos de ellos, pero es posible que no les haya prestado demasiada atención.

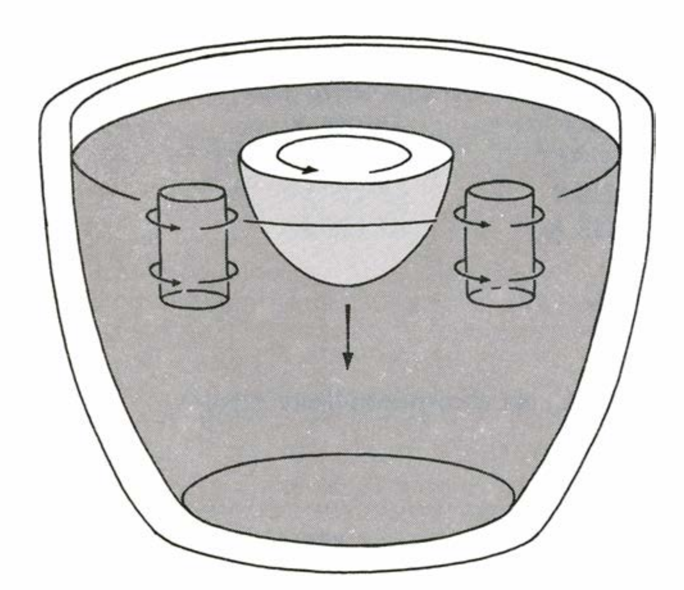
Una noche ya avanzada, mi amigo John Hudak –por evadirse unos minutos del trabajo de su tesis doctoral– estaba distraídamente disolviendo café instantáneo en un tazón de agua caliente; al agitarlo, sonaba un tintineo debido a los choques de su cuchara metálica con el fondo del tazón. El tono del tintineo descendió bruscamente al echar el polvo de café en el agua, lo cual significaba que las ondas sonoras se habían hecho más largas y que había decrecido la frecuencia. Como Hudak continuó un rato removiendo hasta que el polvo se mezcló, la frecuencia incrementó lentamente hasta alcanzar el valor inicial. Obtuvo los mismos resultados cuando añadió una cucharada de un sucedáneo de nata en polvo. En realidad, cualquier polvo disuelto en agua caliente provoca un cambio en la frecuencia del mismo tipo.

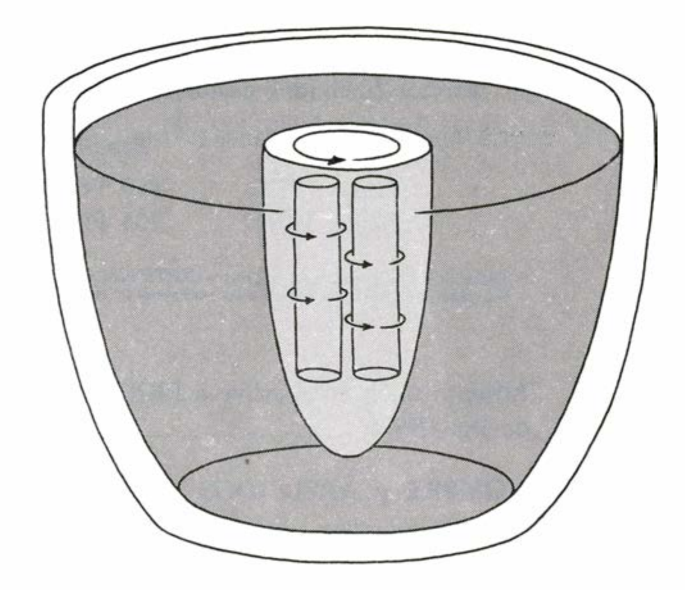
Antes de que se añada ningún tipo de polvo, se oye un sonido peculiar debido al tintineo, ya que la cuchara provoca ondas estacionarias que resuenan a través de la columna de agua y de las paredes del tazón. Las ondas son similares a las que resuenan en un órgano, las cuales son las responsables del sonido que se escucha en el concierto y que sale de los tubos. Sin embargo, los armónicos de frecuencia son un poco más complicados de calcular en una taza de café que en los tubos de un órgano. Las frecuencias dependen del espesor y del radio del tazón, de la densidad y elasticidad del material de que está hecho dicho tazón, del radio de la columna de líquido y de la velocidad de propagación del sonido en dicho líquido.

Lo primero que pensé fue que el polvo, al posarse encima del agua, influiría en la determinación de la frecuencia de resonancia del tazón. Pero resulta que lo único que hace cualquier capa de polvo sobre el agua es amortiguar las oscilaciones en el tazón y, como consecuencia, apagar el ruido del tintineo; no cambia, sin embargo, la frecuencia de resonancia. El polvo se disuelve con cierta rapidez.

Cuando al principio se disuelve el polvo, los gránulos del mismo liberan pequeñas burbujas de aire. La velocidad del sonido es considerablemente menor en el aire (unos 340 metros por segundo) que en el agua (1450 metros por segundo, aproximadamente); las burbujas frenan la velocidad del sonido en el líquido. Si el volumen de las burbujas es la centésima parte del volumen del líquido, la velocidad del sonido se reduce 30 veces. En consecuencia, la frecuencia de cualquier onda estacionaria se hace menor cuando se liberan las burbujas. Como éstas van desapareciendo, gradualmente, la velocidad del sonido (y, por tanto, la frecuencia de la onda estacionaria producida por el tintineo) retorna a su valor normal.

Un cambio similar de frecuencia puede observarse en un vaso de cerveza recién llenado. Conforme las burbujas de aire de la cerveza (no la espuma de encima) se van disipando, el tono del sonido producido por una cuchara al golpear contra la pared del vaso va incrementándose lentamente.





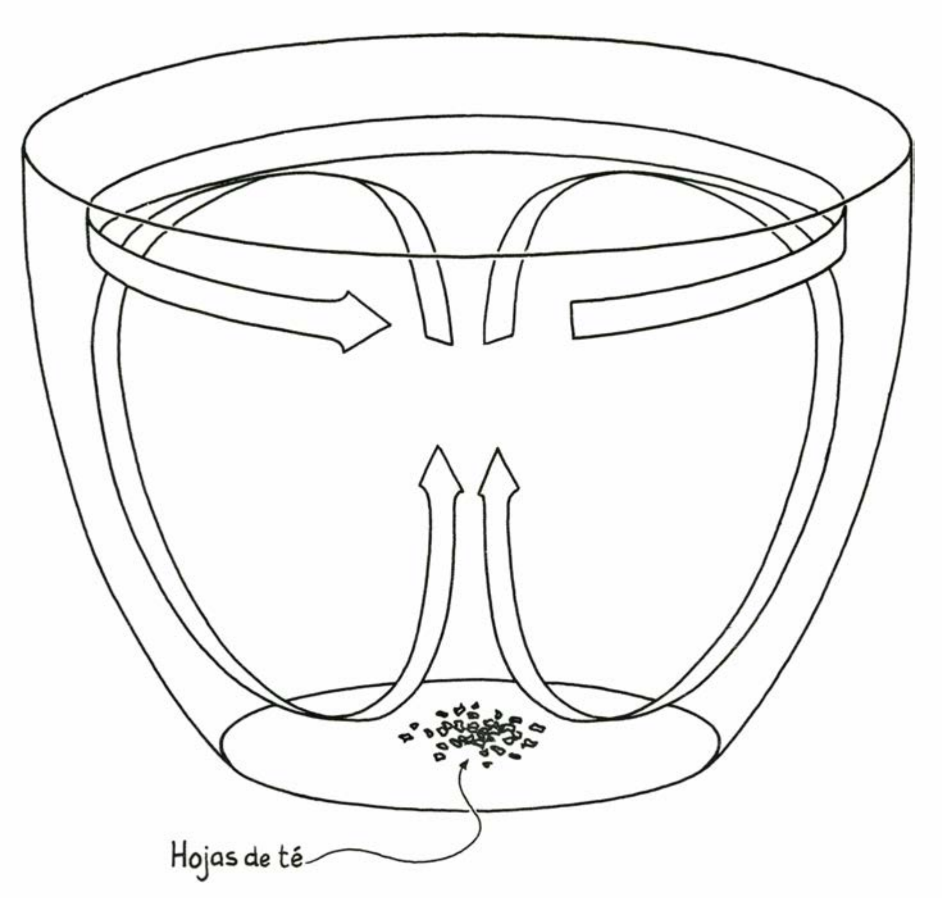
Alargamiento de los torbellinos cuando se vierte leche fría o nata en café caliente

Otro experimento bastante simple con una taza de café sirve como modelo aproximativo de los movimientos de torbellinos o remolinos que tienen lugar a escala mucho mayor en la atmósfera (torbellinos de polvo, tornados, fuegos y huracanes). Remuévase suavemente una taza con café caliente o con té caliente, sáquese la cuchara y a continuación viértase con cuidado leche o nata fría en el centro de la taza. Si el removido y el vertido se han hecho con cuidado, se desarrollará un remolino en el centro del café, el cual se podrá observar por dos razones. En primer lugar, la velocidad angular del fluido es mayor en el centro que justo fuera de la región donde hemos vertido la leche. En segundo lugar, el remolino puede desarrollarse con tanta exactitud que en el centro de la superficie se produzca un hoyito. Estas características no aparecen si la leche está templada o caliente; en ese caso, lo que ocurre es que la velocidad angular en el centro decrece.

Los resultados diferentes que se obtienen al usar leche caliente o leche fría provienen de la diferencia de densidad que ambos líquidos tienen con respecto al agua caliente de la taza. La leche fría es más densa que el agua y por eso se hundirá cuando se eche el agua. Al hundirse, la corriente descendente captura pequeños remolinos verticales que existen en el líquido que la rodea, tirando de ellos hacia el centro y (como término medio) estirándose. La velocidad angular del fluido en la columna aumenta entonces, provocando que el fluido en el centro de la taza haga remolinos más rápidamente que el fluido fuera de la zona ocupada por la leche.

Si la leche está bastante caliente, puede ser más ligera que el agua de la taza. Como no hay corriente descendente (o es mucho más pequeña) los remolinos verticales no se ven arrastrados hacia el centro, y por eso no aumenta la velocidad angular alrededor del centro de la taza.

Puede hacerse una demostración análoga con la leche fría para ampliar remolinos en el aire. El movimiento de convección del aire puede alargar vórtices tubulares ya existentes, exactamente igual a como lo hace la corriente descendente de leche. El alargamiento incrementa la velocidad angular de los remolinos. En el aire, la convección necesaria puede ser hacia arriba (si la corriente es más ligera que el aire que la rodea) o hacia abajo (si la corriente es más pesada que el aire que la rodea). En cualquier caso, la convección puede intensificar los vórtices provocados por otros motivos; ejemplo de ello es el arremolinamiento (debido a un sobrecalentamiento del aire cerca del suelo) que puede dar lugar a torbellinos de polvo.



El flujo secundario en un té removido lleva las hojitas hacia el centro de la taza

Si se agita una taza de café o de té en cuyo fondo se puedan ver unas cuantas partículas, hojitas de té o partículas de azúcar sin disolver, se observará cómo dichas partículas quedan atraídas hacia el centro del fondo. El movimiento resulta un tanto paradójico. Se podría suponer que el material que está en el fondo de la taza debería ser lanzado hacia fuera en dirección a las paredes por la fuerza centrífuga creada por la rotación del fluido alrededor del eje central de la taza.

La fuerza centrífuga intenta empujar el material hacia las paredes. Consideremos una pequeña porción de agua girando alrededor del eje central. Mientras más alejada está dicha porción de agua del eje, mayor es la fuerza centrífuga que ella experimenta. En consecuencia, la presión en el líquido es mayor mientras más alejado esté del eje, con el fin de contrarrestar la fuerza centrífuga. Por tanto, el fluido en su giro no tiene desplazamiento radial sino únicamente desplazamiento angular alrededor del eje.

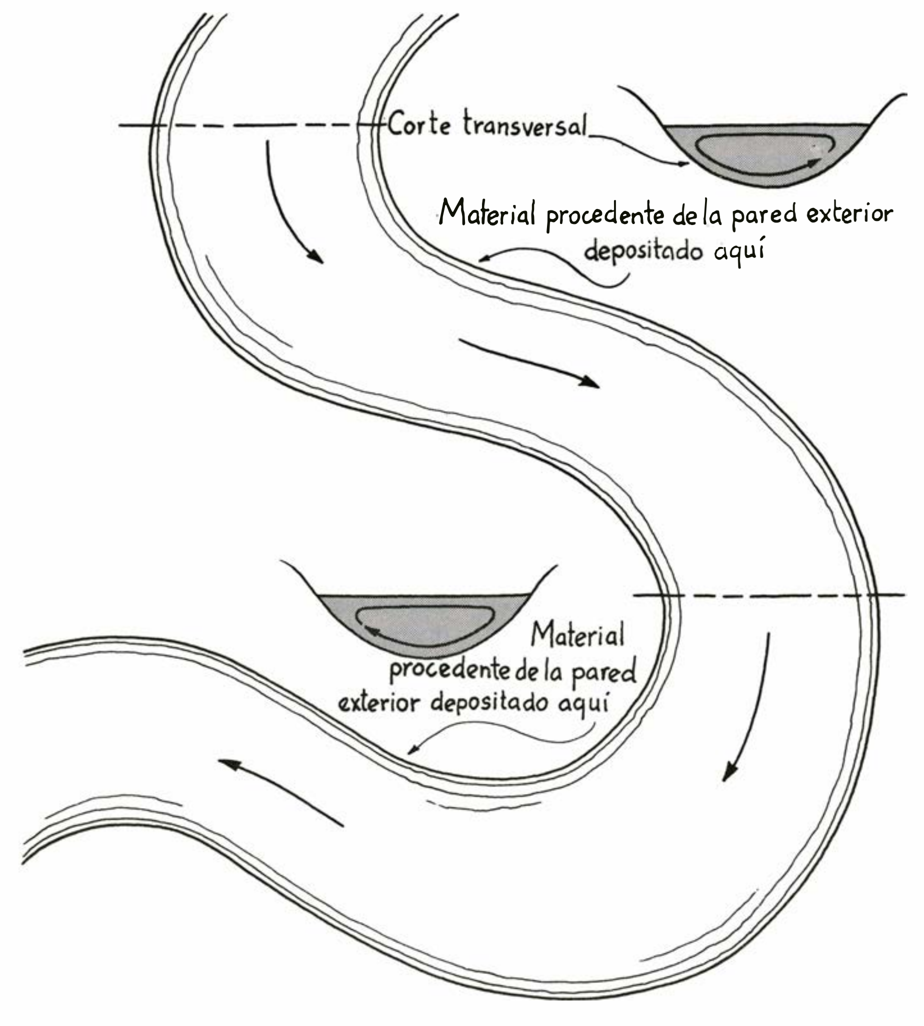
Sin embargo, se produce un movimiento radial debido al rozamiento entre la capa de fluido del fondo y la superficie del fondo de la taza. La reducción de la rotación en la capa del fondo provoca que la diferencia de presión existente entre el fluido que está cerca de la pared y el que está cerca del centro no sea tan grande como la que existe en la capa de arriba. La reducción de presión al girar provoca que la presión cerca de la pared sea mayor en la capa de la superficie que en la capa del fondo. Como consecuencia, el fluido se ve forzado hacia abajo (a lo largo de la pared) y radialmente hacia el interior en dirección al centro de la taza, en el fondo; hacia arriba (a lo largo del eje central) y radialmente hacia fuera, en la capa de la superficie superior. Este movimiento del fluido, conocido como corriente secundaria, puede arrastrar las hojitas de té u otros objetos que estén sumergidos por el fondo de la taza y depositarlos en el centro, al iniciar el fluido su ascensión por dicho centro.

Se puede observar el movimiento de la corriente secundaria haciendo gotear con cuidado un material colorante dentro del agua que se está agitando, en una taza o recipiente transparente. Aunque la corriente del fluido no será tan perfecta como se indica en la figura de esta página, el agua coloreada hará un movimien-

to en espiral alrededor del centro de la taza con otro radial aproximadamente igual al movimiento dibujado.

El mismo tipo de corriente secundaria es responsable en parte de los meandros de los ríos. Similar serpenteo puede verse también en la Corriente del Golfo y en las corrientes de otros océanos, así como en los canales de agua que se forman en la superficie de un glaciar. Aun cuando encontremos una parte de trayecto recta en una corriente, la inspección detenida de la misma nos mostrará que el movimiento del agua no es recto en la realidad, sino que ocurre un serpenteo de un lado a otro entre las paredes de la corriente. Las pequeñas perturbaciones introducidas en el curso de dicha corriente por el terreno que atraviesa inician la formación de los meandros. El misterio radica en la regularidad de las curvas resultantes.

Cuando el agua dobla un recodo, se forma una corriente secundaria similar a la de la taza de té. Debido al retardo de la corriente cerca del fondo, la diferencia de presión entre la parte interior y la parte exterior del recodo es diferente según se trate de una capa de líquido situada en la superficie de la corriente o en el fondo. La mayor presión de la parte exterior de la capa de arriba fuerza al fluido a bajar por la pared exterior. Al ir hacia el fondo, el fluido es impulsado en dirección a la parte interior de la curva, sube luego hacia arriba y, finalmente, sale de nuevo hacia la parte exterior de la curva a lo largo de la capa superficial. Durante ese tiempo, la velocidad horizontal del agua es mayor en la parte exterior de la curva que en la parte interior de la misma. La corriente más rápida arranca troncos de material de la margen exterior; dicho material se ve arrastrado por la corriente secundaria hasta la margen interior. El resultado es un agrandamiento de la curva del cauce, ya que la margen exterior se erosiona y se rellena la margen interior. Si tomamos una corriente que inicialmente empieza recta, la corriente secundaria agrandará cualquier perturbación que haya en su lecho para hacer un pequeño meandro, el cual se irá agrandando hasta convertirse en otro mucho mayor. Si el desarrollo de la curva aumenta notablemente, la propia corriente puede erosionar un nuevo cauce a través de la curva dejando al ensanche aislado al margen.



Efecto del flujo secundario al potenciar la formación de un meandro

Vicent J. Schaefer ha descrito, en *American Scientist*, un sorprendente dibujo geométrico producido sobre la superficie de su café del desayuno. Poniendo la taza de café caliente llena hasta el borde, de manera que la luz brillante del sol la cruce casi horizontalmente, la superficie muestra una zona poligonal como si estuviera llena de polvo, delimitada por trazos negros. Este tipo de dibujos geométricos superficiales fue analizado con detalle por vez primera por Henri Bénard en 1901. Ejemplos como estos dibujos pueden encontrarse en otras evaporaciones de fluidos y sistemas circulantes en la atmósfera y en el mar.

Los dibujos observados en la taza de café se deben a la recirculación del agua caliente que sube del fondo hasta la superficie, se enfría y entonces vuelve a bajar al fondo.

Sobre las zonas en las que el agua caliente asciende, se condensan pequeñas gotas de agua, sostenidas por la presión del vapor existente en la superficie del líquido. Las gotas son de un tamaño bastante uniforme, porque la presión del vapor no puede soportar gotas mayores y si fueran más pequeñas caerían rápidamente. En las zonas donde desciende el agua más fría no existen tales gotas suspendidas y la superficie aparece clara. Al ser oscuro el café, estas zonas aparecen también oscuras. En la superficie pueden observarse pequeñas zonas de gotas (sobre agua ascendente) que están delineadas por unos trazos oscuros relativamente finos donde el líquido desciende.

Si se examinan al microscopio las zonas más ligeras, veremos que están-formadas por capas de gotas muy apretadas entre sí. La densidad de las gotas depende de la presión del vapor que tenga el líquido y del número de núcleos de condensación en el aire. Una atmósfera sucia proporcionará –en principio– más núcleos de condensación que darán lugar a la formación de más gotitas.

Para observar este fenómeno, no se necesita esperar una puesta o salida del sol. Se pueden contemplar más fácilmente los dibujos de recirculación de Bénard dirigiendo la luz de un proyector de diapositivas o de cine sobre la superficie de una taza de café. Si no se quiere gastar café, puede sustituirse por agua caliente teñida con un colorante oscuro (colorante doméstico oscuro o una tinta oscura). Para mantener el calor en el agua puede ponerse la taza (o aún mejor, un recipiente de Pyrex) en una placa caliente a baja temperatura. Para empezar, yo puse

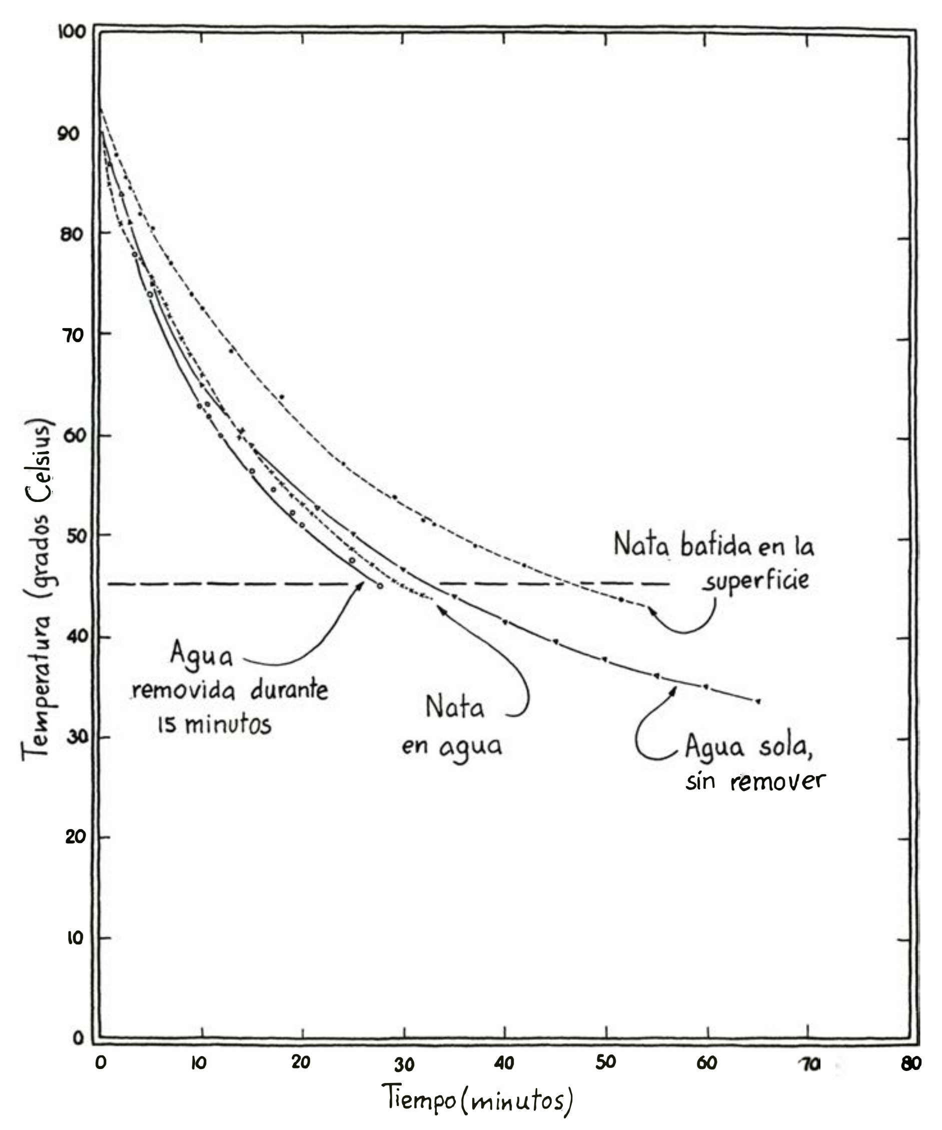
el agua a una temperatura próxima al punto de ebullición para desgasificarla (las burbujas son un estorbo) y después dejé que el agua se fuera enfriando. Las mejores "celdillas" Bénard aparecieron al cabo de varios minutos.

Si se forma una película de suciedad en la superficie, la recojo con una toalla de papel. Si el lector quiere observar el efecto de la contaminación originada por dicha película, puede poner una gota de aceite vegetal en la superficie. Este se extiende lentamente sobre la misma y elimina las "celdillas" Bénard.

Schaefer describió también el efecto causado por la colocación de objetos cargados eléctricamente cerca de zonas de luz en las "celdillas" Bénard. Si hacemos pasar un peine de plástico fuerte por el cabello o a lo largo de unos hilos de lana o de plástico, éste se carga transfiriendo electrones bien al peine o al otro material. Con unos tipos de material, el peine gana electrones y se carga negativamente; con otros, el peine pierde electrones y se carga positivamente. En cualquier caso, el peine cargado destruye las zonas de luz, lo que indica que las gotas ya llevaban una carga.

Cuando Schaefer sostenía un peine de plástico duro cargado cerca de la superficie del agua, vio algunas veces gotas que se formaban cerca de los dientes del peine. Yo he observado ese fenómeno al sostener un cable conductor de un generador electrostático cerca de la superficie, con el otro conductor del generador pegado a un recipiente de metal que contenía agua. En ambos casos, los iones creados en el aire cerca de la punta del objeto cargado son núcleos de condensación para la formación de las gotas. Con una fuente radiactiva ocurre lo mismo porque también produce iones en el aire, exactamente por encima de la superficie del agua. Una cerilla encendida, o cualquier otra fuente de pequeñas partículas, incrementaría de manera similar la formación de gotas.

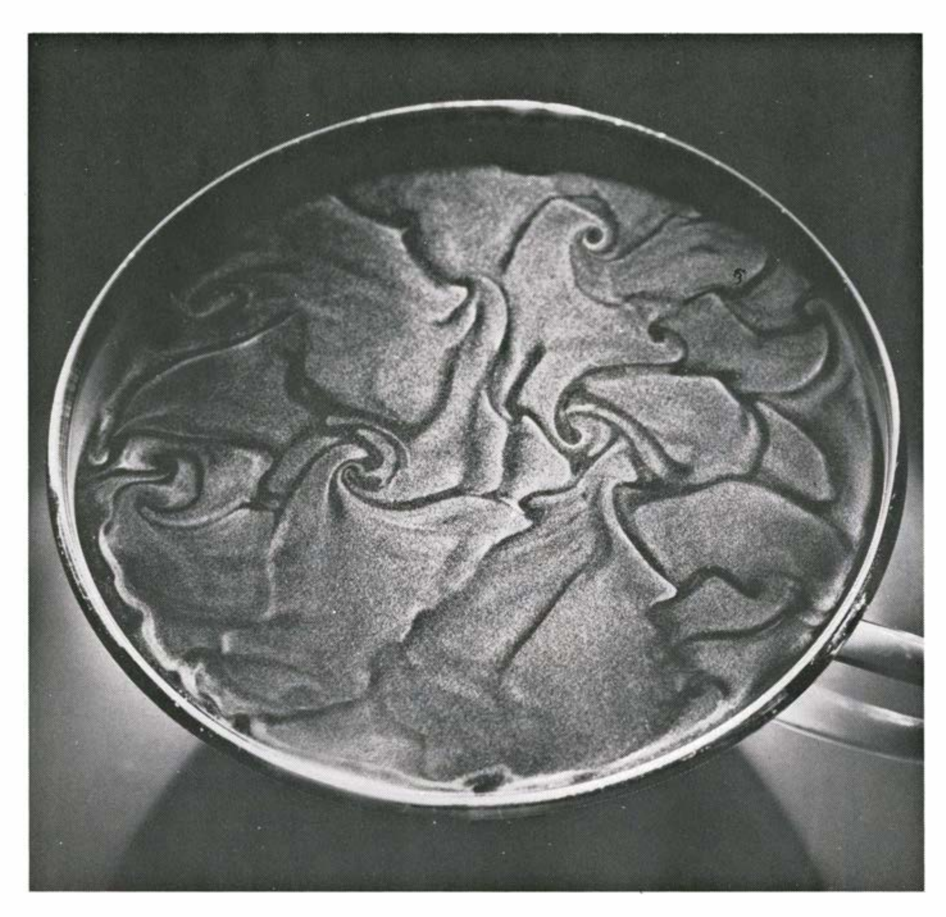
Cuando las gotas están iluminadas por luz blanca fuerte, muestran rápidamente colores cambiantes, delicados e iridescentes. La dispersión de luz que da origen a estos colores se reconoce por dispersión de orden superior de Tyndall, en honor de John Tyndall quien empezó a investigar sobre la misma en 1869. La dispersión de Tyndall puede observarse también en otros fenómenos naturales. La corona que se ve a veces alrededor del sol, que puede mostrar varios espectros completos de colores, resulta de la dispersión de Tyndall de pequeñas gotas en las finas nubes que se interponen entre la persona que mira el sol y éste (no nos referimos aquí a la corona que puede contemplarse durante el eclipse). El mismo origen tiene la corona que se observa alrededor de la luna. La dispersión causa también los delicados colores que aparecen en las nubes de color madreperla o nacaradas que se contemplan a veces después de la puesta del sol en las altas latitudes. Los modelos matemáticos de esta dispersión son complejos, porque el tamaño de las gotas se aproxima a las longitudes de onda de la luz visible. Las gotas miden aproximadamente una micra de diámetro y, por tanto, están situadas entre gotas mayores que pueden producir arcos iris (dispersión geométrica) y las partículas más pequeñas que dan origen al color azul del cielo (dispersión de Rayleigh).



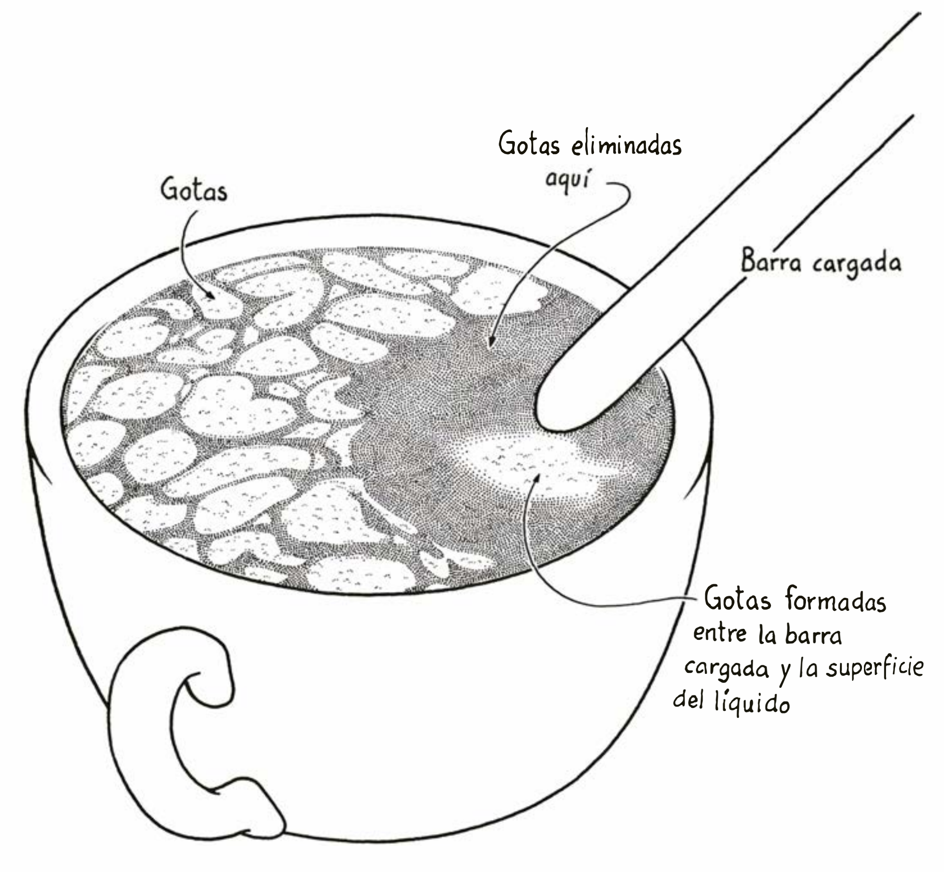
Velocidad de enfriamiento del agua caliente

Además de los cambios constantes en la superficie de la taza de café o cualquier otro recipiente, el lector encontrará también líneas oscuras que atraviesan rápidamente las zonas que parecen tener polvo. Schaefer afirma que estas líneas se deben a pequeños remolinos que se desarrollan justo sobre la superficie y duran sólo una fracción de segundo. La capa de aire que está inmediatamente encima del líquido es inestable porque está más caliente que el aire situado ligeramente más arriba. El lector puede crear su propio vórtice sujetando una ficha bibliográfica en sentido vertical con un borde cerca de la superficie. Cualquier pequeña corriente de aire que sople más allá del borde provoca un vórtice que se desliza a lo largo de la superficie. Incluso sin el borde, se desarrollan a veces remolinos mayores sobre la superficie, que llegan a durar unos 10 segundos mientras transportan rápidamente el vapor y las gotitas del entorno.

¿A qué velocidad se enfría una taza de café o de té? Para hacer retroceder las fronteras de la física moderna (por lo menos un milímetro o dos) medí la velocidad a la cual el agua se enfriaba de la temperatura a la que se había calentado



Dibujos geométricos sobre la superficie de un café caliente



Efecto causado por la colocación de una barra cargada cerca de la superficie

para preparar un café instantáneo. Quería saber cuál de los distintos procedimientos haría que el agua se enfriara más rápidamente. Herví agua en una tetera eléctrica y eché 200 mililitros en un recipiente de Pyrex con capacidad para 250 mililitros, el cual dejé sobre una placa de metal. Medí la temperatura del agua con un termómetro (con escala de 0 a 100 grados Celsius); lo introduje en el recipiente con el depósito de mercurio apoyado en una esquina del fondo del recipiente. La temperatura inicial en todas las pruebas fue 93 grados Celsius. Saqué mediciones de la temperatura hasta que el agua estaba por debajo de 45 grados, temperatura que creía adecuada para que una taza de café no resultara desagradablemente fría.

El agua sin aditivos y sin moverla se enfriaba suavemente, alcanzando 45 grados en 33 minutos, aproximadamente. Poniendo la medida equivalente a una cucharadita de café instantáneo en el recipiente antes de echar el agua, ésta seguía el mismo ritmo de enfriamiento durante los primeros 15 minutos, pero después se iba enfriando más rápidamente de lo que lo había hecho en el primer ensayo. Los resultados fueron casi idénticos cuando se pusieron tres terrones de azúcar en el recipiente y cuando se dejó una cuchara de metal en el mismo (sin añadir nada más). Se podría pensar que la inclusión de café instantáneo en polvo enfriaría el agua más rápidamente por dos razones: al estar el polvo a temperatura ambiente, absorbe calor del agua y obscurece también la disolución, incrementando por tanto la radiación térmica de alguna manera. El último efecto debe ser despreciado, dado que la adición de azúcar al agua no obscureció la solución, pero causó efectos similares. De la cuchara se puede esperar que incremente la velocidad de enfriamiento del agua porque inicialmente absorbe parte del calor del agua y después actúa como un radiador de calor en una habitación.

En otra prueba añadí 20 mililitros de nata líquida a una temperatura inicial de 10 grados Celsius, inmediatamente después de haber echado el agua. Revolví el agua sólo tres veces para mezclar la crema sin enfriarse por causa de la remoción. La temperatura descendió inmediatamente unos 4 grados, pero después de cinco minutos la mezcla de nata y agua empezó a seguir la misma curva de enfriamiento que había seguido el agua sola; sin embargo, después de 15 minutos, la mezcla de nata se empezó a enfriar más rápidamente.

La adición de alcohol podría incrementar la velocidad de enfriamiento de-

bido a la rápida evaporación de la superficie. Después de haber añadido 20 mililitros de vodka de 80 grados, a temperatura ambiente, la velocidad de enfriamiento era casi la misma que cuando añadí nata.

De todos los aditivos, el más eficaz para la alteración de la curva de enfriamiento del agua fue la nata Reddi (una nata batida que se presenta enlatada a presión) la cual eché en la superficie del agua, como se hace con el café irlandés. La crema estaba fría y pudo haber hecho descender inicialmente la temperatura de la superficie del agua, pero su efecto principal fue el de retener el calor en el agua y eliminar la evaporación. Ello comportó que el agua tardase 14 minutos más en alcanzar los 45 grados Celsius.

Me imaginé que, moviendo el agua tal cual, incrementaría mucho la velocidad de enfriamiento. Sin embargo, el efecto de agitar con fuerza el agua mediante el termómetro durante los 15 primeros minutos, redujo la temperatura del agua en 2 grados solamente. Al parecer, la remoción que yo realizaba no era mucho mejor que la convección normal de las moléculas en el agua al transportar agua caliente desde el centro a la superficie.

En mi último ensayo cubrí de negro las paredes exteriores del recipiente. Con una superficie negra, las paredes deberían radiar calor mejor, haciendo que el líquido se enfriara más de prisa. El agua se enfrió antes, repitiendo casi la curva de enfriamiento obtenida al removerla.

De todos estos datos se desprenden dos conclusiones. Primero, si se quiere enfriar el café rápidamente, pero sin añadirle mucha nata ni mucho azúcar, lo más indicado es ponerlo en una taza negra y removerlo bien con una cuchara de metal. No obstante, el enfriamiento resultante apenas merece tal esfuerzo, ya que sólo rebaja dos grados. Segundo, si se desea mantener el café lo más caliente posible, lo mejor que puede hacerse es servir un café irlandés y relajarse.

Si se estuviera interesado en seguir trabajando sobre este problema podrían considerarse los efectos resultantes de añadir nata, azúcar o un sucedáneo de nata en polvo, a los cinco o diez minutos de haber echado el agua. ¿Desciende la temperatura del café si se permite que éste se enfríe por evaporación y convección antes de añadir la nata, el sucedáneo en polvo o el azúcar? Puede que le guste descubrir la dependencia que existe entre el enfriamiento del café y la graduación de alcohol del whisky en un café irlandés. La investigación sobre este asunto podría ser muy atractiva si el experimento se lleva a cabo entre dos.

Libros

La obsesión por el movimiento continuo y un completo manual sobre los adhesivos

Philip Morrison

PERPETUAL MOTION: THE HISTORY OF AN OBSESSION, por Arthur W. J. G. Ord-Hume. St. Martin's Press. Hacia la época de las catedrales la idea ya se había asentado: un "número impar" de mazos colgando de una rueda podían hacerla girar por sí sola. Alrededor de 1640, el marqués de Worcester probablemente mostró una rueda semejante "de catorce pies y pico", y que llevaba sobre palancas 40 pesas de 50 libras, ante Carlos I en la Torre de Londres. (La inercia es impresionante, y los volantes eran casi desconocidos.) Hizo el marqués una breve relación de la rueda en su *Century of Inventions* ("Siglo de invenciones"), bajo el inexpresivo encabezamiento de *An Advantageous Change of Centres* ("Un ventajoso cambio de centros"). Se dibujó, y a veces se construyó, molinos de agua ideados para bombear agua hacia arriba y dejarla caer sobre la rueda, así como un ciclo de molino de viento-fuelle. Hubo los escépticos de siempre.

El Obispo Wilkins escribía en 1648 que "apenas si pude abstenerme de gritar, con Arquímedes, '¡Eureka!'", cuando tuvo por primera vez la idea de la rueda hidráulica, pero la experiencia y el ensayo le convencieron de que la "corriente, aunque multiplicada, no tendrá la fuerza suficiente para hacer girar un tornillo". Simon Stevinus usó la cadena desequilibrada de bolas que corren en torno a un plano inclinado para deducir el equilibrio de fuerzas no paralelas invocando la imposibilidad del movimiento continuo. "Una maravilla, pero que no es de maravillar", escribió. Sin embargo, en 1730, los volúmenes más famosos sobre mecanismos seguían mostrando las máquinas cíclicas perpetuas, y empezaron a aparecer fraudes sobre ellas en diversos recintos reales. Continuó la avalancha de esperanzadas patentes. La primera patente inglesa de un mecanismo de movimiento continuo se concedió en 1635, y se emitieron 600 o más hasta 1903. La mayoría pertenecieron a la época victoriana, y el tema fue analizado en las páginas escépticas de *Scientific American* que estaban prendadas del auge de los inventores norteamericanos.

Los mecanismos llegaron a perder su misterio, e innovadores tales como lo fueron en su tiempo Jean Bernoulli, Robert Boyle y William Congreve empezaron a invocar la capilaridad (Congreve dispuso pesadas cadenas que presionaban de modo desigual sobre esponjas empapadas), la asimetría de flotabilidad, el magnetismo, y, con el tiempo, motores y generadores eléctricos cuyos extremos estaban conectados entre sí.

Se abre la sesión de fraudes públicos. Filadelfia albergó dos de ellos. El primero, hacia 1813, fue exhibido por un tal Charles Redheffer. "Irrefragablemente un principio automotriz perpetuo", desplegaba un gran engranaje propulsado por un piñón pequeño. Un despierto jovenzuelo observó que estos engranajes se habían gastado por la cara incorrecta: la potencia estaba proviniendo a través del supuesto eje pasivo. Redheffer llevó su artefacto a Nueva York (pero sólo después de que los ingeniosos ciudadanos de Filadelfia se hubiesen resarcido y le hubiesen mostrado su máquina de movimiento continuo, propia de ellos. Impulsada por un muelle de reloj oculto, sigue siendo una obra maestra de despistar bobos, con pequeñas pesas que se deslizan sobre planos inclinados, dispuestas de modo que la relojería interior no podía impulsar las ruedas a menos que las pesas estuviesen presentes). En Nueva York, fue el formidable Robert Fulton quien exclamó al entrar en la sala: "¡Pero si esto es un movimiento de manivela!". Advirtió el ritmo de una mano invisible, inesperado en motores continuos, y echó abajo unos livianos tirantes de madera que parecían servir para apuntalar la máquina. Quedó al descubierto una trasmisión de tripa que condujo hasta "un viejo de barba larga", que hacía girar la manivela motriz en un cuarto del altillo.

Unos 50 años más tarde, John W. Keely, empresario de Filadelfia, de cabeza fría pero escasa educación, vendió acciones de su "motor hidro-neumático-pulsante al vacío" a miles de especuladores de la Edad de Oro, que recibieron una hermosa ilustración del mismo en sus bonos. Dicho motor extraía "fuerza etérica" de una cucharadita de agua. La firma obtuvo millones de dólares, pero no vendió ni un julio de energía. Después de ocho o 10 años de vida por todo lo alto, Keely quedó sin blanca. El relato periodístico de su lucha con la pobreza atrajo la atención de una mujer, viuda reciente de marido rico, que se convirtió en su benefactora. La máquina de Keely revivió, "vibratoria" esta vez, y calculada para llevar un barco a través del Atlántico con un par de litros de agua. El tema siguió en discusión hasta 1896, a pesar de que unos escépticos habían advertido cables huecos –que se presentaban como sólidos– en casi cada una de las piezas del aparato del laboratorio de Keely. Eran abastecimientos de aire comprimido, supuso la gente, y así resultó ser. Después de la muerte de Keely, un examen de su casa reveló conductos de aire, imanes móviles, ejes y fuentes de presión hidráulica, todo ello escondido en los muros e impulsado por un fuerte y silencioso motor a cuerda colocado en el sótano.

Existe hoy en el sur de California un proyecto bien financiado de ingeniería química para liberar hidrógeno combustible del agua sin costos de energía. El hidrógeno está allí, por supuesto; pero todos aguardamos los resultados.

El autoengaño es más simpáticamente ingenuo que el fraude. En 1918, una comisión de expertos creada por ley del Congreso emprendió el estudio oficial del proyecto de un hombre honrado, Garabed Giragossian. Este había conmovido al país entero, y en especial a la prensa, con su pública oferta: pedía sólo un minucioso escrutinio de expertos cuando llegara el momento en que ofreciera al público de los Estados Unidos su nuevo sistema de obtener energía gratis: ¡por fin el movimiento continuo! Se había dirigido directamente al Congreso porque dudaba en la protección de la Oficina de Patentes. ¡Su invento era el volante! Vio que podía hacer andar su rueda, una vez puesta en marcha, con un motor de 1/20 de caballo de fuerza y que, cada vez que se la detenía bruscamente, producía 10 caballos de fuerza. Totalmente honrado, jamás ha-

bía entendido la diferencia entre energía y potencia.

El libro anota algunas interesantes y caprichosas digresiones. Un reloj hecho en la década 1760-69 estaba movido por las variaciones de la presión atmosférica, la cual inducía movimientos pequeños, pero fiables, de la columna de un barómetro para subir así las pesas. El Movimiento Continuo de Mr. Cox, un bello centro de interés en su tiempo, funcionó bien usando la abundancia de la atmósfera durante toda una vida; hoy descansa, mudo, en el Victoria and Albert Museum, con su pesada carga de mercurio extraída años ha, cual noble ruina. Ord-Hume puede consolarse algo con el hecho de que una conocida marca suiza fabrica hoy en día una línea de hermosos y prácticos relojes de sobremesa con la misma virtud de no necesitar que les den cuerda, pero con una cámara aneroide en vez de una columna de mercurio. En un laboratorio de Oxford hay actualmente un carillón eléctrico movido por dos pilas voltaicas secas, plantadas en azufre en 1840. La diminuta corriente fluye aún para hacer sonar las campanas unas cuantas veces por segundo, excepto cuando la humedad es excesiva. El badajo se está desgastando visiblemente, y bajo él yace "un minúsculo pero ominoso montón de polvo metálico". ¿No fue éste acaso el predecesor de la moderna circuitería integrada de baja potencia?

Hay buen humor, historia generalmente cuidadosa, y sentido de la física en este bien ilustrado libro no técnico, hecho por un ingeniero aeronáutico inglés. No todos sus análisis son muy pertinentes, pero abundan los excelentes. Es difícil olvidar el encuentro del propio Ord-Hume, en 1958, con un conmovedor y anciano londinense, que en su pequeño apartamento había construido un nuevo y maravilloso motor de aviación que funcionaba sin fallos a gran potencia, sin combustible ninguno. El autor, naturalmente, preguntó: "¿Podría verlo?". La cortés respuesta fue aún más razonable. Por supuesto había estado funcionando allí mismo, allí en el vestíbulo, durante semanas, pero el inventor no podía hacerlo parar, de modo que los demás inquilinos habían insistido en que lo desarmara por completo ¡para terminar con el ruido, tan molesto!

HANDBOOK OF ADHESIVES, manual dirigido por Irving Skeist. Van Nostrand Reinhold Company. "En un mundo de ominosa desorganización, los adhesivos se dedican a combinar y unir". Realizan esa sutil tarea bidimensional suya, más o menos visiblemente, a través de toda la tecnología contemporánea, desde el tablero de madera contrachapada o de yeso de nuestras paredes hasta el calzado de nuestros pies, las cabezas de las cerillas de nuestros bolsillos, los libros en rústica de nuestras estanterías, los sobres y sellos que lamemos, las bolsas y cajas de supermercados que llevamos, y la moqueta de pared a pared que cruzamos en el vestíbulo. En Estados Unidos se producen cuatro millones de toneladas anuales. Aunque la mayor parte son polímeros sintéticos, incluyendo una amplia variedad de cauchos, para muchos usos la substancia pegajosa proviene de huesos de vacuno, sangre de cerdo, pieles de pescado, caseína de leche descremada y tratada con ácido, almidón de maíz y de sorgo y fibras de residuo de algodón; todos ellos, polímeros naturales.

Este robusto volumen, en segunda edición, es un interesante representante de un conocido género. En él, unos 75 especialistas unen los actuales 56 capítulos en torno a la línea de encolado, esa fina frontera de la adhesión. Los primeros capítulos, que han de atraer al lector general, pasan revista a la teoría, a la importancia económica, a la especificación y a la prueba de los adhesivos; el grueso del libro trata, un tipo químico tras otro, desde el vidrio soluble de antaño hasta la goma de bibliotecario, la cola blanca del colegial, una amplia variedad de sintéticos, las últimas siliconas y los polímeros orgánicos aún más recientes, que sirven como pegalotodo al rojo fuego. El volumen termina con capítulos (de ingeniería mecánica más bien que química) que repasan los usos y técnicas, incluyendo las cintas sensibles a la presión, los abrasivos aglutinados y los pegamentos y cementos de la práctica médica y odontológica. (El adhesivo proteínico natural del disco de sujeción del mejillón común se queda pegado, mojado y seco, a pesar de los embates oceánicos, a "todas las superficies conocidas", incluidos los teflones.) Es obvio el valor del volumen como referencia, aunque se mueve en un nivel técnico elevado. Lo que le presta valor es la abundancia con que trasmite mediante ejemplos, mejor que cualquier artículo de encuesta afirmativo, el ingenio y el vigor de una tecnología creciente y diversa.

La teoría de los adhesivos se remonta a Thomas Young y a Willard Gibbs, quienes dieron tratamientos simples pero fundamentales, mecánicos y termodiná-

micos respectivamente, al esparcimiento y a la capilaridad. La unión adhesiva es, ante todo, un problema de interacción líquido-sólido. En cada junta hay al menos cinco componentes: las dos tajadas adherentes, el relleno adhesivo del emparedado y las dos interfaces de "mantequilla". El primer acto es la mojadura de las superficies sólidas por el adhesivo, complicada como toda física de superficies por la aspereza de las superficies, vapores adsorbidos, esfuerzos internos de las superficies e interacción química entre la superficie sólida y líquida. Una generación de cuidadosos estudios realizados en torno a W. A. Zisman, del Laboratorio de Investigaciones Navales de los Estados Unidos, ha comenzado a poner sobre bases cuantitativas la química de la elección de adhesivos. Los antiguos conceptos termodinámicos sirven solamente si está controlada la naturaleza de las superficies.

La superficie se moja y el líquido se esparce libremente siempre que la energía de superficie del líquido sea menor que la del sólido en las condiciones reales. A temperaturas corrientes, la mayoría de los líquidos tienen energías de superficie del mismo orden de magnitud que la del agua, y en los sólidos cristalinos duros y de elevado grado de fusión dichas energías son hasta cien veces mayores. Sin embargo, los polímeros orgánicos (papel, madera, caucho) que nos agrada pegar son sólidos de escasa energía de superficie, que en ellos tiene valores bastante parecidos a los de los líquidos, por lo que el no esparcimiento es frecuente.

El progreso principal ha consistido en un conjunto de cuidadosas mediciones que establecen, para muchos tipos de sólidos, cuál es la energía de superficie crítica para la mojadura. Partiendo de una limpieza extremada, materiales puros, superficies lisas (como el vidrio) y conocimientos químicos prácticos sobre el efecto de los vapores contaminantes, los ángulos de contacto de las gotas del líquido son reproducibles, con independencia de que la gota esté avanzando o retrocediendo sobre la superficie, y son químicamente razonables. Una monocapa de moléculas adsorbidas –de agua, por ejemplo– esconderá en gran parte la mayoría de las superficies sólidas respecto de los fluidos; en el caso de sustancias no polares, las fuerzas son las fuerzas de dispersión de Fritz London, que decaen rápidamente al aumentar la distancia; en el de las sustancias polares, las fuerzas dipolares de los enlaces de hidrógeno constituyen un mecanismo molecular aún más fuerte.

Una controversia reciente veló los verdaderos cimientos del tema por más de cinco años. ¿Moja el agua libremente las superficies de oro puro pulido, con el ángulo nulo de contacto exigido por los conceptos termodinámicos? Hacia la presente década la respuesta era claramente afirmativa. Esta química de muestras de materiales a granel que se disponen efectivamente en capas de sólo algunas moléculas de espesor es el triunfo del virtuosismo racional en la técnica, salvando la fe que teníamos en los conceptos simples.

Finalmente, la unión pegada fragua (salvo en las cintas adhesivas descascarables que todos usamos). Aquí rige la física del estado sólido. El encogimiento y la flexión originan tensiones. Una junta encolada suele desunirse cuando se le somete a la décima parte del esfuerzo calculado según las energías de superficie.

Estamos en el dominio de la propagación de grietas. Las asperezas superficiales tienden a permitir la formación de burbujas de aire en la interfaz de "mantequilla" del antedicho emparedado, y los esfuerzos concentrados propagan una grieta de rotura a lo largo de un plano de burbujas como un cierre de cremallera. Lo convencional es hacer rugosa la superficie encolada, lo cual aumenta el área de contacto, pero las burbujas coplanares constituyen una clara indicación en sentido contrario. La aspereza es un concepto muy burdo; lo que cuenta es la forma de la superficie. La adherencia indeseada se rige por la misma teoría y, tecnológicamente, tiene asimismo gran importancia. En todo moldeo, fundición y laminado, la adherencia es un problema. Piénsese además en la escarcha en alas y parabrisas. Quizás una aspereza del tipo adecuado pudiese debilitar la garra helada, pero no es fácil. Cuando congelación y deshielo se repiten, el hielo puede adherirse incluso al teflón, pues se mete en los poros y adquiere una gran área superficial de contacto.

Vale la pena relatar ciertas muestras de ingenio. El papel engomado se encrespa al secar, como pronto aprende el empapelador de paredes aficionado. Tradicionalmente se alisa para la impresora "quebrándolo" sobre el agudo borde de una regla de acero. Los diminutos segmentos de película quebrada de goma pueden dilatarse o contraerse sin ejercer una gran tensión sobre el papel. A los impresores se les facilita la tarea median-

te adhesivos inescrespantes que no requieren el quiebre mecánico de la goma y, sin embargo, permanecen planos dentro de una gama más amplia de humedades. Se micropulveriza partículas discretas de cola, se forma con ellas una suspensión fangosa en un solvente volátil sin agua y se lamina esta suspensión sobre una cinta o rótulo. Luego se extrae el solvente para usarlo de nuevo. Las fibras de papel se mojan por primera vez al pegarlas en un lugar.

Hay adhesivos anaerobios que reemplazan a las arandelas de seguridad y a otros accesorios en muchos conjuntos mecánicos. El oxígeno inhibe la polimerización del monómero; hasta el aire que se difunde hacia el interior de una botella de polietileno mantiene líquido el compuesto en la estantería durante un año, o más. Pero basta con extender la substancia aislada sobre un perno, y atornillar éste en su rosca, para la exclusión del oxígeno permita el fraguado tenaz tan pronto como se agote el oxígeno residual. Los iones de hierro catalizan el proceso de manera que el endurecimiento es rápido, sin liberación de agua, de solvente, o de calor siquiera, muy dentro de un altavoz o de un motor eléctrico. La radiación ultravioleta endurece en unos segundos ciertos adhesivos.

Los alquilcianoacrilatos forman rápidamente fuertes enlaces con casi todos los materiales. Su polimerización viene iniciada por las sustancias alcalinas, y hasta el vapor de agua es lo bastante básico como para hacerlo. Una gota (el pegamento se vende a más de 200 dólares el kilogramo) puesta entre dos superficies limpias de acero lijado de una pulgada cuadrada (645,16 milímetros cuadrados) las unirá de modo que resistan un esfuerzo cortante de más de una tonelada. Los alquilcianoacrilatos se usan mucho en cirugía, principalmente como suplementos para las suturas.

Se ha creado un ingenioso aparato carente de piezas móviles para mezclar íntimamente dos componentes fluidos. En un tramo de tubo se disponen una serie de tapones idénticos de aluminio hábilmente perforados. Dos corrientes líquidas, que penetran en el primer elemento, pasan a través de esas complejas canalizaciones para emerger como ocho filetes fluidos reordenados en cuanto a posición. Un conjunto de 10 de estos elementos, llamados generadores de superficie interfacial, corta y baraja un par de millones de capas de fluido.

El libro contiene un glosario y un índice completísimo; la primera edición apareció en 1962.

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

LA BUSQUEDA DE VIDA EN MARTE

SPECIAL VIKING ISSUE. *Science,* vol. 193, n.º 4255; 27 de agosto de 1976.

SPECIAL VIKING ISSUE. *Science,* vol. 194, n.º 4260; 1 de octubre de 1976.

SPECIAL VIKING ISSUE. *Science,* vol. 194, n.º 4271; 17 de diciembre de 1976.

SPECIAL VIKING ISSUE. *Journal of Geophysical Research,* vol. 82, n.º 28; 30 de septiembre de 1977.

RIEGO GOTA A GOTA

DRIP IRRIGATION: PRINCIPLES, DESIGN, AND AGRICULTURAL PRACTICES. Dan Goldberg, Baruch Gornat y Daniel Rimon. Drip Irrigation Scientific Publications, Israel, 1976.

WELD CONTROL UNDER DRIP IRRIGATION IN ORCHARD AND VINEYARD CROPS. A. Lange, H. Kempen, B. Fisher, F. Aljibury y J. Schlesselman en *Proceedings of the 2nd International Drip Irrigation Congress, San Diego, California,* págs. 422-424.

EL AGRUPAMIENTO DE LAS GALAXIAS

NEW REDUCTION OF THE LICK CATALOG OF GALAXIES. Michael Seldner, B. Siebers, Edward J. Groth y P. James E. Peebles en *The Astronomical Journal,* vol. 82, n.º 4, págs. 249-256; abril, 1977.

STATISTICAL ANALYSIS OF CATALOGS OF EXTRAGALACTIC OBJECTS, VII: TWO- AND THREE-POINT CORRELATION FUNCTIONS FOR THE HIGH RESOLUTION SHANE-WIRTANEN CATALOG OF GALAXIES. Edward J. Groth y P. James E. Peebles en *The Astrophysical Journal,* vol. 217, parte 1.ª, págs. 385-405; 15 de octubre de 1977.

GATOS Y VIAS COMERCIALES

COMPARATIVE GENETICS OF COAT COLOUR IN MAMMALS. Anthony G. Searle. Academic Press, Inc., 1968.

MELANIN, ADRENALIN AND THE LEGACY OF FEAR. Clyde Keeler, Theodore Meliinger, Edward Fromm y Lenore Wade en *The Journal of Heredity,* vol. 61, n.º 2, págs. 81-88; marzo-abril, 1970.

THE EFFECTS OF SELECTION AND HUMAN PREFERENCE ON COAT COLOUR GENE FREQUENCIES IN URBAN CATS. J. M. Clark en *Heredity,* vol. 35, parte 2.ª, págs. 195-210; octubre, 1975.

GENETICS FOR CAT BREEDERS. R. Robinson. Pergamon Press, 1977.

LOS USOS DE LOS INSTRUMENTOS DE SILEX DEL PALEOLITICO

PREHISTORIC TECHNOLOGY: AN EXPERIMENTAL STUDY OF THE OLDEST TOOLS AND ARTEFACTS FROM TRACES OF MANUFACTURE AND WEAR. S. A. Semenov, traducción e introducción de M. W. Thompson. Barnes & Noble, Inc., 1964.

MICROWEAR ANALYSIS OF EXPERIMENTAL FLINT TOOLS: A TEST CASE. Lawrence H. Keeley y Mark H. Newcomer en *Journal of Archaeological Science,* vol. 4, n.º 1, págs. 29-62; marzo, 1977.

EL PROCESO DE LA FECUNDACION

ACTIVATION OF SEA-URCHIN EGGS BY A CALCIUM IONOPHORE. Richard A. Steinhardt y David Epel en *Proceedings of the National Academy of Sciences,* vol. 71, n.º 5. págs. 1915-1919; mayo, 1974.

THE ISOLATION OF INTACT CORTICAL GRANULES FROM SEA URCHIN EGGS: CALCIUM IONS TRIGGER GRANULE DISCHARGE. Victor D. Vacquier, en *Developmental Biology,* vol. 43, págs. 62-74; 1975.

THE SURFACE EVENTS AT FERTILIZATION: THE MOVEMENTS OF THE SPERMATOZOON THROUGH THE SEA URCHIN EGG SURFACE AND THE ROLES OF THE SURFACE LAYERS. Gerald P. Schatten y Daniel Mazia en *Journal of Supramolecular Structure,* vol. 5, n.º 3, págs. 343-369; 1976.

INTRACELLULAR *p*H AND ACTIVATION OF SEA URCHIN EGGS AFTER FERTILIZATION. James D. Johnson, David Epel y Miles R., Paul en *Nature,* vol. 262, n.º 5570. págs. 661-664; agosto 19, 1976.

CONSECUENCIAS DE UNA ANTERIOR CRISIS ENERGETICA

NOTE ON THE PROGRESS OF IRON PRODUCTION IN ENGLAND, 1540-1640. John U. Nef en *The Journal of Political Economy,* vol. 44, n.º 3, págs. 398-403; junio, 1936.

SILVER PRODUCTION IN CENTRAL EUROPE, 1450-1618. John U. Nef en *The Journal of Political Economy,* vol. 49, n.º 4, págs. 575-591; agosto, 1941.

THE AGRICULTURAL REVOLUTION. Eric Kerridge. Allen & Unwin Ltd, 1967.

WAR AND HUMAN PROGRESS: AN ESSAY ON THE RISE OF INDUSTRIAL CIVILIZATION, John U. Nef. W. W. Norton & Co., 1968.

LA ESTRUCTURA DE LOS CONCEPTOS CIENTIFICOS

FUNDAMENTALS OF CONCEPT FORMATION IN EMPIRICAL SCIENCE. Carl G. Hempel. Chicago University Press. Chicago, 1952.

THEORIE UND ERFAHRUNG. Wolfgang Stegmüller. Springer-Verlag. Heidelberg, 1970.

THEORY OF MEASUREMENT. J. Pfanzagl. Physica-Verlag. Würzburg-Wien, 1971.

FOUNDATIONS OF MEASUREMENT. D. H. Krantz, R. D. Luce, P. Suppes y A. Tversky. Academic Press. New York-London, 1971.

JUEGOS MATEMATICOS

ON EDGEWISE 2-COLORED GRAPHS WITH MONOCHROMATIC TRIANGLES AND CONTAINING NO COMPLETE HEXAGON. Ronald L. Graham en *Journal of Combinatorial Theory,* vol. 4, n.º 3, pág. 300; abril, 1968.

GRAPH THEORY. Frank Harary. Addison-Wesley Publishing Company, Inc., 1969.

GENERALIZED RAMSEY THEORY FOR GRAPHS–A SURVEY. Stefan A. Burr en *Lecture Notes in Mathematics, Vol. 406, Graphs and Combinatories: Proceedings of the Capital Conference on Graph Theory and Combinatorics at the George Washington University, June 18-22, 1973,* págs. 52-75; 1974.

TALLER Y LABORATORIO

ON THE NOTE EMITTED FROM A MUG WHILE MIXING INSTANT COFFEE. W. E. Farrell, D. P. McKenzie y R. L. Parker in *Proceedings of the Cambridge Philosophical Society,* vol. 65, Parte 1.ª, págs. 365-367; enero, 1969.

OBSERVATIONS OF AN EARLY MORNING CUP OF COFFEE. Vincent J. Schaefer en *American Scientist,* vol. 59, n.º 5, págs. 534-535; septiembre-octubre, 1971.